

Potential salinity and temperature futures for the Chesapeake Bay using a statistical downscaling spatial disaggregation framework

Barbara A. Muhling^{1,2*}, Carlos F. Gaitán³, Charles A. Stock², Vincent S. Saba⁴, Desiree Tommasi^{1,2}, Keith W. Dixon²

1: Princeton University Program in Atmospheric and Oceanic Science, Forrestal Campus/Sayre Hall, Princeton, NJ 08544, USA.

2: NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ 08540, USA

3: Arable Labs Inc., Princeton, NJ 08542, USA

4: National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, Geophysical Fluid Dynamics Laboratory, 201 Forrestal Road, Princeton University Forrestal Campus, Princeton, NJ 08540, USA

* Barbara.Muhling@noaa.gov

Abstract

Estuaries are productive and ecologically important ecosystems, incorporating environmental drivers from watersheds, rivers, and the coastal ocean. Climate change has potential to modify the physical properties of estuaries, with impacts on resident organisms. However, projections from general circulation models (GCMs) are generally too coarse to resolve important estuarine processes. Here we statistically downscaled near-surface air temperature and precipitation projections to the scale of the Chesapeake Bay watershed and estuary. These variables were linked to Susquehanna River streamflow using a water balance model, and finally to spatially-resolved Chesapeake Bay surface temperature and salinity using statistical model trees. The low computational cost of this approach allowed rapid assessment of projected changes from four GCMs spanning a range of potential futures under a high CO₂ emissions scenario, for four different downscaling methods. Choice of GCM contributed strongly to the spread in projections, but choice of downscaling method was also influential in the warmest models. Models projected a ~2 - 5.5°C increase in surface water temperatures in the Chesapeake Bay by the end of the century. Projections of salinity were more uncertain and spatially complex. Models showing increases in winter-spring streamflow generated freshening in the upper Bay and tributaries, while models with decreased streamflow produced salinity increases. Changes to the Chesapeake Bay environment have implications for fish and invertebrate habitats, as well as migration, spawning phenology, recruitment, and occurrence of pathogens. Our results underline a potentially expanded role of statistical downscaling to complement dynamical approaches in assessing climate change impacts in dynamically challenging estuaries.

Keywords: Chesapeake Bay, statistical downscaling, spatial disaggregation, climate change

Introduction

Estuaries are highly productive coastal environments, which provide a broad range of ecosystem services to human communities (Beck et al. 2001; Martinez et al. 2007). They supply essential nursery habitats for several diadromous fishes, and support commercial and recreational fisheries targeting multiple species (Costa et al. 2002; Pihl et al. 2002). In addition, estuarine environments are important for nutrient cycling and filtration processes, and also function as natural storm protection (Martinez et al. 2007; Barbier et al. 2011). However, due to their accessibility to human populations, estuaries are often subject to heavy anthropogenic pressures (Edgar et al. 2000). They form a focal interface between terrestrial, river and ocean environments, and tend to concentrate and retain nutrients and pollutants (Barbier et al. 2011; Feyrer et al. 2015). Along with substantial landscape and flow regime modification, this has led to the eutrophication and degradation of ecosystem function in many estuaries worldwide (Cloern et al. 2016).

In addition to these management challenges, there is increasing recognition of the potential for climate change impacts on estuaries. As estuaries are transitional systems between land and ocean, they will incorporate complex drivers from both environments, across multiple temporal and spatial scales (Scavia et al. 2002; Gillanders et al. 2011; Cloern et al. 2016). Changes to temperature and precipitation cycles in watersheds, including snow melt dynamics, will affect freshwater inflow to estuaries, shifts in flow peaks, extreme events, and droughts (Gleick and Adams 2000; Milly et al. 2005; Vigano et al. 2015; Demaria et al. 2016; Lee et al. 2016). Along with land use patterns, changes to flow cycles may modify nutrient and sediment inputs to estuaries, affecting eutrophication, habitats, and ability to meet management targets for water quality (Howarth et al. 2002; Scavia et al. 2002; Chen et al. 2014). Rising sea levels will result in increased coastal inundation, and altered circulation patterns and flushing characteristics (Scavia et al. 2002). Warming air temperatures will drive increases in estuarine water temperatures, with impacts on physiological stress, phenology, migration, and recruitment of estuarine-dependent species (Hare and Able 2007; Wagner et al. 2011; Bell et al. 2014; Peer and Miller 2015; Wilber et al. 2016). Many diadromous fishes are already endangered, threatened or vulnerable as a result of habitat loss, overfishing, and pollution (Jelks et al. 2008; Limburg and Waldman 2009). Their reliance on estuarine habitats will likely render them highly vulnerable to the effects of climate change (Lynch et al. 2014; Hare et al. 2016).

There is clearly a need to develop and refine projections of future conditions in estuaries for the purposes of planning and management. However, current general circulation models (GCMs) have typical spatial resolutions of >100km. This is too coarse to resolve processes at the scale of most estuaries, or even the scale of most watersheds (Stock et al. 2011). Thus, GCMs are generally downscaled to a more appropriate spatial scale, using dynamical or statistical downscaling methods (Wood et al. 2004; Wagner et al. 2011; Flato et al. 2013). Statistical downscaling works by assuming that the local-scale climate is a product of both large-scale climatic processes, and smaller-scale local processes. Relationships between these can then be used to develop future projections of local conditions (e.g. Cannon and Whitfield 2002; Johnson and Weaver 2009; Gaitán et al. 2014; Dixon et al. 2016). The advantages of statistical downscaling are primarily that it is computationally simpler and faster than dynamical downscaling, and that it incorporates bias-correction inherently. A key feature is that the computational simplicity allows for rapid generation of an ensemble of projections spanning a range of climate futures, GCMs, greenhouse scenarios, and internal climate variations (Hawkins and Sutton 2011). Incomplete exploration of this ensemble of climate futures is a primary limitation of the past generation of climate impacts assessments on marine resources (Cheung et al. 2016; Payne et al. 2016). The main disadvantages of statistical downscaling are that stationarity is assumed (i.e. that the relationships between regional and local-scale processes remain constant as climate changes), and that long historical observational datasets are required (Wilby and Wigley 1997; Diaz-Nieto and Wilby 2005; Benestad et al. 2008; Dixon et al. 2016;). There is also a wide range of different downscaling methods to choose from, further expanding the range of climate futures that can be considered (Hessami et al. 2008; Chen et al. 2013; Gaitán et al. 2014).

Statistical downscaling is likely to be well suited for estuarine environments, as many have comprehensive time series data available for key variables such as temperature and salinity (Feyrer et al. 2015; Cloern et al. 2016; Schulte et al. 2016). However, many modeling steps are required to get from coarse resolution atmospheric outputs from GCMs to projections of biologically relevant environmental conditions in the estuary itself. There are multiple choices to be made regarding input variables and model structure, with complex and interacting sources of uncertainty. Different studies have approached these questions differently: using one to many GCMs and a range of downscaling methods and hydrological models, across estuaries with different hydrological and environmental characteristics (e.g. Maurer and Duffy 2005; Wilby and Harris 2006; Vicuna and Dracup 2007; Chen et al. 2014; Thompson et al. 2015; Brown et al. 2016).

The Chesapeake Bay (**Fig. 1**) is the largest estuary in the United States, supporting multiple biological communities, ecosystems, and human use activities, and providing essential habitat for a large number of economically important fish and invertebrate species (Richards and Rago 1999; Sharov et al. 2003; Najjar et al. 2010). This importance has motivated decades of physical and biological observations to track and understand the response of the Chesapeake Bay to changes in land use, climate, and other potential stressors (Hagy et al. 2004; Kimmel and Roman 2004; Kemp et al. 2005; Najjar et al. 2010). Water quality in the Chesapeake Bay is strongly tied to the timing and magnitude of freshwater inflow events, through influences on nutrient and sediment delivery, flushing times, and stratification (Gibson and Najjar 2000; Glibert et al. 2001; Wood et al. 2002; Paerl 2006; Paerl and Otten 2013; Lee et al. 2016). Historical land use practices in the watershed have led to significant eutrophication of the Bay, with consequent declines in water clarity, submerged aquatic vegetation, and bottom oxygen concentrations in warmer months (Sprague et al. 2000; Boesch et al. 2001; Langland et al. 2004; Kemp et al. 2005). Harmful algal blooms have also been increasing in incidence and severity, as a consequence of both nutrient enrichment and warming waters (Glibert et al. 2001; Paerl and Otten 2013). These events can contribute to hypoxia, fish kills, and seafood contamination (Tango et al. 2005). Water temperature in the Bay can also drive phenology and recruitment of various fish species (Hare and Able 2007; Bell et al. 2014), while streamflow may trigger migration behaviors of diadromous species (Tommasi et al. 2015). There is therefore a substantial need to assess potential climate change impacts on streamflow, temperature, and salinity regimes in the Chesapeake Bay, and how these may interact with current management issues.

While the data richness of Chesapeake Bay is conducive to statistical downscaling, the dynamical complexity poses challenges. The Bay is nearly 300 km long, and up to 48 km wide. While it has a relatively deep (>20 m) and narrow central channel with reasonably well characterized dynamics, much of the Bay consists of shallow and often complex habitat, with an average depth of only 6.5 m across the entire estuary (Hagy et al. 2005). More than half of the freshwater inflow comes from the Susquehanna River, which drains a 71,250 km² watershed stretching across urban, suburban, and rural areas of the states of Maryland, Pennsylvania and New York (Schubel and Pritchard 1986; Hagy et al. 2005). This primary watershed combines with numerous smaller ones to shape salinity and circulation conditions across the Bay (Guo and Levinson 2007; Shen and Wang 2007; Reay and Moore 2009; Xu et al. 2012). While up-estuary portions of the Chesapeake Bay are strongly shaped by these river flows, down-

estuary conditions reflect a mix of river input and continental shelf conditions. These dynamical complexities may impact the feasibility of statistical downscaling approaches.

In this study, we developed a simple modeling framework to obtain projections of surface temperature and salinity in the Chesapeake Bay from multiple GCMs, downscaled using different statistical downscaling techniques. A key novel feature of this framework for the Chesapeake Bay is the ability to assess spatial structure of surface temperature and salinity. Our approach attempts to draw a balance between capturing the primary estuarine responses to climate drivers, while maintaining the computational efficiency required to rapidly assess a range of climate futures. We assess the sufficiency of this approach in capturing past variations in the surface hydrography of the Chesapeake Bay, and then examine primary drivers of estuarine conditions, and contributions to uncertainty from spread in GCM projections vs. downscaling methods. Potential consequences of both future change, and the uncertainty around these projections, are then discussed for the Chesapeake Bay ecosystem.

Methods

A conceptual diagram of the overall modeling framework is shown in **Fig. 2a**. Daily air temperature and total precipitation from GCMs were statistically downscaled, and fed into a simple water balance model to derive Susquehanna River streamflow. Daily air temperature was also used to estimate daily surface water temperature at the Thomas Point buoy (**Fig. 1**). Streamflow and water temperature were then combined with other variables to project spatial patterns of surface temperature and salinity across the Chesapeake Bay, at monthly resolution. Each step of this process is described in more detail below.

Thomas Point Water Temperature Model

Before the full framework could be used to generate future projections for the Chesapeake Bay, the development of several models relating air temperature and rainfall to water temperature and salinity using historical observations was required. The first and simplest of these predicted daily surface water temperature from daily air temperature (**Fig. 2a**). The best long-term time series of air and water temperature in the Chesapeake Bay is from the Thomas Point buoy, which has been recording since fall 1985 (**Fig. 1**). Several studies (e.g. Hare and Able 2007; Hare et al. 2010; Tommasi et al. 2015) have used daily air temperature as a direct proxy for water temperature in estuaries and rivers, due to their low heat capacity relative to deeper ocean waters, often yielding a high correlation between the two variables. However, there are some issues with this approach. Firstly, the relationship between air and

water temperature is often non-linear, leveling off at very cold and very warm temperatures (Mohseni et al. 2003). As climate change projections will require extrapolation of present-day relationships, behavior at extremes is important to quantify. Leveling off at temperature extremes was clearly evident at Thomas Point. The daily air and water temperatures were strongly correlated ($R^2 = 0.89$, RMSE = 3.56°C), but the relationship was obviously non-linear, particularly at cold temperatures. Differences between daily air and water temperatures could exceed 5°C at cooler times of year (**Fig. 3**).

Water temperatures integrate air temperatures over the preceding days to weeks, and may therefore lag them considerably, particularly during times of rapid air temperature change (Letcher et al. 2016). To account for this, we determined the optimum integration time for air temperature to predict daily water temperature. Correlations were calculated between daily water temperature (1985 – 2015) and moving means of daily air temperature, the latter tested at all values between 2 and 21 days. Of all the values tested, a 17 day moving mean best improved the correlation ($R^2 = 0.98$, RMSE = 3.08°C) (**Fig. 3**).

Once the lag issue had been addressed, the problem of non-linearity was solved using the exponential equation from Mohseni et al. (2003):

$$T_w = \mu + \frac{\alpha - \mu}{1 + e^{y(\beta - T_a)}}$$

This equation was optimized using the non-linear least squares (nls) function in R 3.2.1 (R Core Team, 2015). The best fit between 17 day air temperatures and water temperatures was achieved where $\mu = -6.53$, $\alpha = 37.08$, $y = 0.1$ and $\beta = 14.49$. While this equation was initially designed for unsmoothed daily data, it was also useful for further improving the fit of the 17 day air temperature vs. daily water temperature model, particularly at cooler temperatures ($R^2 = 0.99$, RMSE = 3.04°C) (**Fig. 3**). This improvement in correlation was statistically significant at $p < 0.001$ (Fisher r-to-z transformation).

Water balance model

A wide array of hydrological models are available for estimating streamflow from temperature, precipitation and other variables. We chose to test the sufficiency of a simple water balance model to maximize efficient and rapid consideration of multiple potential climate futures. Our water balance model requires only air temperature and precipitation, and runs at a monthly resolution, using Java code

provided by McCabe and Markstrom (2007) adapted into R 3.2.1. The model assigns precipitation in the watershed to snow or rain, depending on temperature. Results are aggregated for the whole watershed, and no routing models are included. Snow is stored in snowpack, with melt rates determined by temperature, while rain and snowmelt contribute to runoff after soil moisture storage is saturated. Runoff across the entire watershed area then becomes streamflow at the mouth of the Susquehanna River. Actual evapotranspiration is determined from potential evapotranspiration (PET), soil moisture storage, and soil moisture storage withdrawal. Of the many methods available for calculating PET, we chose to use the Hamon equation, which has previously shown low error and bias in U.S. watersheds (Vorosmarty et al. 1998). All model parameters were set at the values recommended by McCabe and Markstrom (2007), except for the maximum snowmelt proportion, which was set to 0.7 instead of 0.5 (within the range of parameter uncertainty). This adjustment gave predictions slightly closer to observations. Extractive water use in the catchment is currently <5% (Najjar 1999; SRBC 2013), and so we did not consider this in the water balance model.

Historical precipitation and air temperature were obtained for the Susquehanna River watershed from the NOAA/NCEP GPCP 0.5° monthly temperature dataset (CAMS: Fan and van den Dool 2008), and the CPC Unified Gauge-Based 0.25° Analysis of Daily Precipitation (CPC: Chen et al. 2008). All grid points inside the watershed were averaged by year and month, from 1970 to 2006, and used to drive the water balance model. Results were compared to monthly streamflow observations at Conowingo Dam, which is ~16 km north of where the Susquehanna River opens into the northern end of the Chesapeake Bay (USGS station 01578310, obtained from 1970 to 2006 from http://waterdata.usgs.gov/nwis/dv/?referred_module=sw). Although the Potomac, James, and other rivers also deliver freshwater into the Chesapeake Bay, time series of monthly flow from all major rivers were highly correlated (e.g. Susquehanna vs. Potomac: $R^2=0.65$, 1970 - 2006, USGS). As a result, we only modeled flow from the Susquehanna River for our framework, to maximize simplicity and minimize multicollinearity in predictive models.

While modeled monthly streamflow from the water balance model (1970 - 2006) was well correlated with observed streamflow at Conowingo Dam ($R^2 = 0.76$), predictions were biased low, particularly during winter (January – March) and spring (April – June) (**Fig. 4**). This suggested a problem with snow measurements, which are well known to be affected by wind-induced under-catch and bias from the placement of gauges (Hayhoe et al. 2007). Larsen and Peck (1974), and Adam and Lettenmaier (2003)

suggested that snow under-catch varied strongly by region, but averaged around 50% at wind speeds typical of the Susquehanna watershed in winter-spring ($4 - 7 \text{ ms}^{-1}$: NCEP/NCAR Monthly Reanalysis). Iterative testing of a snow catch ratio in the water balance model showed that a value of 0.55 was optimal (i.e. precipitation classified as snow by the water balance model should be divided by 0.55). This simple correction improved the bias of the model substantially, and also slightly improved the fit ($R^2 = 0.8$) (**Fig. 4**). However, very high flow events (e.g. Hurricane Agnes, 1972) were still somewhat underestimated by the model. Note that the snow correction was not required for projections from the GCMs, as it relates only to bias in actual precipitation observations from gauges.

Spatial temperature and salinity models

Spatial models were built using historical CTD cast data from the Chesapeake Bay Program (<http://www.chesapeakebay.net/data>), the University of Maryland Chesapeake Biological Laboratory cruise database (hjordt.cbl.umces.edu), and the Smithsonian Environmental Research Center database (<https://serc.si.edu/environmental-data>), from 1986 - 2015. Surface temperature and salinity values were extracted from individual casts at all stations within the Chesapeake Bay and major tributaries. Stations <0.5 km from shore were excluded. This represented a small overall proportion of the pelagic, sub-tidal Chesapeake Bay, but excluded sampling stations subject to finer-scale nearshore variability, which our models would likely not be able to capture (e.g. Breitburg 1990). The CTD data set was used to estimate spatially-resolved surface temperature and salinity within the Chesapeake Bay, given estimates of Thomas Point surface temperature and streamflow (derived as described above) and a small set of additional predictors described below.

Daily mean near-surface air temperature at Thomas Point was strongly correlated to air temperature at ten other nearby buoys in the Chesapeake Bay ($75.8 - 77^\circ\text{W}$ and $37 - 39.5^\circ\text{N}$) for years 2010 - 2014 ($R^2 = 0.83$ to 0.93), confirming highly coherent daily air temperature variation across the Chesapeake Bay. The estimated Thomas Point buoy temperature thus served as the primary predictor of surface water temperature at other latitudes and longitudes via spatial covariance. We also included three other predictors supported on both mechanistic and theoretical grounds: a) the 30 day trend in Thomas Point temperature change to account for seasonal hysteresis in spatial temperature covariance (Letcher et al., 2016), b) freshwater inflow, which can have a pronounced cooling effect in the upper portions of the Bay relative to moderating ocean effects in its lower reaches (Preston, 2004), and c) the time of day to account for the diurnal signal in CTD casts (**Table 1**). Variables a) and b) are point measurements, taken

at only one spatial location (i.e. Thomas Point and the Conowingo Dam, respectively). We therefore included latitude and longitude as predictors in the model, to allow the modeling of surface temperature and salinity in two dimensional space. The same predictor variables were used for the spatial model of surface salinity, except that predicted Thomas Point temperature was not included.

The Cubist package in R 3.2.1 (Kuhn et al. 2015) was used to create statistical model trees to predict surface temperature and salinity across the Chesapeake Bay (**Fig. 2b**). These models are similar to standard regression tree models, in that they split the training data into increasingly similar subsets based on the values of predictor variables, before arriving at predicted values at terminal nodes. Model trees differ, however, in that the values at the terminal nodes are described using multivariate linear equations (terminal linear models), rather than fixed values. These linear equations predict the value of the target outcome (in this case, temperature or salinity), based on a subset of the predictor variables. The model tree is thus reduced to a set of conditional linear regression equations or “rules”, which can then be either eliminated via pruning, or combined for simplification (Kuhn et al. 2015). As a result of their use of linear equations at tree nodes, model trees can extrapolate beyond the range of training datasets, while many other machine learning techniques cannot (Quinlan 1992, 1993). This characteristic is particularly important for modeling temperature under future climate change, as warming will lead to novel conditions outside the range of historical datasets. Model tree training can be refined using “committees”, which operate similarly to boosting for boosted regression trees (Elith et al., 2008). Essentially, multiple model trees are constructed, each one learning from the deficiencies of the previous one, and the final predicted value is a mean from all trees (Kuhn et al. 2015).

Machine learning predictive models are flexible and powerful, but can be prone to overfitting if appropriate steps are not taken to control this. They are generally expected to perform much better on training rather than unseen test or validation datasets (e.g. Elith et al. 2008), with high skill on training data and low skill on validation data indicative of overfitting, and potentially an overly complex model. To determine the best model configurations, model trees were therefore trained on the first 20 years of CTD data (1986 – 2005), and validated on the last 10 years (2006 – 2015). The optimal number of control rules and committees was determined by assessing root mean square error (RMSE) only on the unseen test data. The model configuration that gave the best results on the test data was thus considered to be sensitive enough to capture important relationships and interactions, but general enough to avoid overfitting. The optimum configuration to predict surface temperature was 100 committees, and a

maximum of 20 rules, while for surface salinity it was 100 committees, and a maximum of 15 rules. The importance of the predictor variables to both the conditional splits of the model tree, and the linear models at the terminal nodes, were reported as percentages, as described in Kuhn et al. (2015). The maximum importance score that a predictor can attain is 100%. Thus the total percentage score across all variables will not add up to 100%.

As surface temperature and salinity both have very strong spatial characteristics in the Chesapeake Bay, the predictive power of the two model trees was assessed across 7 major sub-regions: the James, Rappahannock, Potomac and Patuxent Rivers, and the Upper, Mid and Lower portions of the Bay main-stem (**Fig. 1**). The Upper Bay was defined as all main-stem waters north of the Patuxent River mouth, the Lower Bay was defined as waters south of the Rappahannock River mouth, and the Mid Bay was all main-stem waters in between. In addition, surface temperature results were assessed using monthly anomalies, to remove the effect of the very strong seasonal temperature signal (i.e. the observed monthly mean surface temperature for each sub-region was subtracted from both observed and modeled values before comparison).

Selecting GCMs

Our aim was to consider a range of plausible futures for the Chesapeake Bay, using GCMs from the 5th Coupled Model Intercomparison Project (CMIP5). All GCMs (n = 33) available which included 2m air temperature and total precipitation were assessed for inclusion. Model output was accessed through the NOAA Climate Change Web Portal (Scott et al. 2016), which re-grids GCMs to a common spatial resolution, and models were compared for the Chesapeake Bay and Susquehanna River watershed region (36 - 43°N, 74 - 80°W). Late 20th century (1956 - 2005) climatologies of air temperature and precipitation from GCMs were compared to observations from the CAMS temperature dataset, and the CPC precipitation analysis for the same region and time period. As model hindcast skill within a region of interest often differs among variables, selecting a subset of “best models” can be difficult (Overland et al. 2011; Sheffield et al. 2013). We instead focused on culling models with the highest hindcast error, and then selecting models that encompassed the range of future temperature and precipitation projections for use in our study. Outlier models were defined as those with an annual air temperature error of >2°C, and/or a precipitation error of >400 mm year⁻¹. Seven of the 33 candidate GCMs were removed using these criteria: FGOALS-S2, MIROC-ESM, MIROC-ESM-CHEM, CAN-ESM2 and ACCESS1-3 were excluded for warm temperature bias (>2°C) versus observations. FGOALS-G2 was too cool, and

rainfall in CMCC-CESM was too high. Adjusting total observed rainfall upwards with the snow correction described above did not change the model selections.

The spread of future projections from the remaining models highlighted a broad range of potential futures for the region under RCP8.5 (**Fig. 5**). All GCMs projected warming temperatures, and some increase in mean annual precipitation in the Susquehanna River watershed. The extent of these trends, however, varied considerably among models. We chose four GCMs to capture the range of potential futures: GFDL-CM3 (more warming, greater precipitation increase: hereafter WW (warm wet) model), IPSL-CM5A-LR (more warming, less precipitation increase: hereafter WD (warm dry) model), GFDL-ESM2G (less warming, less precipitation increase: hereafter CD (cool dry) model), and MRI-CGCM-3 (less warming, greater precipitation increase: hereafter CW (cool wet) model) (**Fig. 5**). Each of these GCMs had a spatial resolution of >90 km over the study area. While INMCM4 and FIO-ESM were both more extreme examples of cooler, drier GCMs than GFDL-ESM2G, they had very strong seasonal bias compared to historical observations (winter several degrees too warm, summer in INMCM4 also too cold). This bias persisted after downscaling, and so we selected ESM2G instead.

Daily air temperature and precipitation were extracted from each GCM, at all grid points contained within the Susquehanna River watershed (**Fig. 1**). In addition, air temperature was extracted for the closest grid point to the Thomas Point buoy. Where several grid points were close to this location, the one where the mean late 20th century seasonal air temperature cycle most closely resembled that observed at Thomas Point (1985 - 2000) was chosen.

Statistical Downscaling

Statistical downscaling was used to derive a) estimates of Thomas Point air temperature, and b) estimates of air temperature and precipitation across the Chesapeake Bay watershed from coarse resolution GCM data. We note that the latter case involves deriving anomalies across a scale that can include several GCM grid cells but which nonetheless remains poorly resolved by GCMs.

Four statistical downscaling methods were applied to outputs from the 4 GCMs: Bias-Corrected Quantile Mapping (BCQM), Change Factor Quantile Mapping (CFQM), Equidistant Quantile Mapping (EDQM), and the Cumulative Distribution Function Transform (CDFt). These statistical models are all based on quantile mapping but differ in their implementation of the bias-correction step. We acknowledge that

there are more sophisticated ways in which to derive quantile mapping functions (e.g. Cannon 2011) and that there are many more statistical downscaling techniques that also could have been used (e.g. Zorita and von Storch 1999; Wilby et al. 2002; 2003; Fasbender and Ouarda 2010; Zeng et al. 2011; Gaitán et al. 2014). However, we chose to include the selected methods, which are conceptually simple and relatively low-cost, to demonstrate the flexibility of our modeling framework. Our approach also investigates whether the downscaled projections from these related methods may differ, with implications for the use of multiple downscaling methods for impact assessment (Gaitán 2016).

Of the aforementioned quantile methods, the most commonly employed is the BCQM (Ho et al. 2012). It derives separate bias corrections at each position (quantile) within the cumulative distribution function (CDF). In essence, the bias, defined as the difference between the observations and climate model during the historical period, is used as a correction factor for the model output during the future. The second method, the CFQM, (see Ho et al. 2012 for details), is much less commonly used in practice. In essence, the change in the model from the historical to the future period is applied to the historical observations, again separately at each position in the CDF.

Since there is no obvious theoretical argument to favor either of the previous two strategies, Li et al. (2010) introduced the Equidistant Cumulative Distribution Function matching method (EDQM), combining aspects of the CFQM and BCQM methodologies. A criticism of the BCQM method is that it assumes that the bias computed for the historical period is applicable to the future (Li et al. 2010). Conversely one could criticize the CFQM approach since it assumes that the change factor computed for the climate model is equally applicable to the observations. The EDQM method deals with these issues by applying a correction that consists of two terms: one for bias correction and the other as a change factor.

The fourth method, Cumulative Distribution Function transform (CDFt; Michelangeli et al. 2009), uses the following function to obtain the cumulative distribution function of the downscaled variable:

$$F_{\text{obsF}}(x) = F_{\text{obsH}}(F_{\text{gcmH}}^{-1}(F_{\text{gcmF}}(x))),$$

This method uses a series of post-processing refinements to improve the tail behavior of the projections (unlike the raw BCQM method, described above). Specifically, the time-series from the coarse-resolution GCMs are transformed to have the same mean as the local historical time series, thus preventing the

downscaled projections getting out of range when implementing the transform equation. Hence, the final output maintains the initial mean from the corresponding GCM output (historical or future).

Specifically, we used the CDFt R package (Vrac and Michelangeli 2009) to obtain the downscaled time series. This implementation needs two parameters to be defined: *npas* and *dev*. *Dev*, or the coefficient of development, is used to extend the range of data on which the quantiles will be calculated, while *npas* is the number of quantiles to be empirically estimated (default value is 100). Here, we used *dev* = 1 and *npas* = default.

Each grid point in the Susquehanna River watershed from the CPC precipitation analysis ($n = 119$, 1970 - 2005) was assigned to the closest grid point for each GCM ($n = 2 - 6$: **Fig. 1**). A mean of all CPC grid points assigned to each GCM point was then used as the historical precipitation observations for downscaling. Daily air temperature observations were obtained from eight weather stations for 1970 - 2005. Similarly to precipitation, each weather station was assigned to the closest grid point for each GCM, and a mean taken to provide historical observations air temperature observations for downscaling. Air temperatures at Thomas Point were obtained from only one grid point from each GCM, and were downscaled using historical observed air temperatures at the Thomas Point buoy (1985 – 2015).

Temperature and precipitation outputs from statistically downscaled GCMs were then run through the framework shown in **Fig. 2a**, to give projections of surface temperature and salinity across the Chesapeake Bay. As the water balance model ran at monthly resolution, all predictor variables input to the model trees were also aggregated to month and year before this analysis. Results thus provided projections of surface temperature and salinity at monthly resolution. Model outputs were compared between the late 20th century (1970 - 1999) and the late 21st century (2071 - 2100), to show possible temperature and salinity futures from different GCMs.

Results

Spatial surface temperature and salinity estimates

Both the surface temperature and surface salinity models reproduced historical spatiotemporal variability across the Chesapeake Bay reasonably well, despite the dynamical complexity of the system. The predictive power of the surface temperature model on monthly anomalies (using only the out-of-sample test years 2006-2015) was highest in the Lower Bay (i.e. most seaward: $R^2=0.74$) and James River ($R^2 = 0.77$), and lowest in the Rappahannock River ($R^2 = 0.62$) (**Table 2**). RMSE averaged 1.32°C , and was

highest in the Rappahannock River (1.51°C), and lowest in the Lower Bay (1.14°C). The most important variables for the conditional splits in the model tree were the modeled water temperature at Thomas Point (75%), and the seasonal (30 day) change in air temperature (71%). These two variables were also most important to the terminal linear models, with scores of 100% and 76%, respectively.

The surface salinity model gave the best results in the Upper Bay ($R^2 = 0.76$), and was weakest in the Lower Bay and the Patuxent River ($R^2 = 0.55$, $R^2 = 0.42$, respectively) (**Table 2**). RMSE averaged 2.21, and was highest in the James River (3.00), and lowest in the Upper Bay (1.80). Absolute RMSE in the Potomac River was comparable to other regions (2.02), but constituted the highest percentage of observed mean annual salinity (49.5%) (**Table 2**). The most important variables for the conditional splits in the model tree were latitude (97%), longitude (90%) and streamflow (41%). These three variables were also most influential for the terminal linear models, with scores of 97%, 93% and 97%, respectively.

Example time series of monthly observed and modeled surface temperature (anomalies) and surface salinity for the Upper Bay, Lower Bay and Potomac River show that the model trees generally tracked observed values well (**Fig. 6**). As expected, model skill was high during the 1986 – 2005 training period, and degraded somewhat in the unseen 2006 – 2015 testing period. However, R^2 statistics between observed and predicted surface temperature anomalies stayed above 0.6 for the test period in all zones of the Chesapeake Bay. Monthly time series of surface salinity in the test period were also reasonably well represented in the James River, Mid Bay, Potomac River, Rappahannock River, and Upper Bay ($R^2 > 0.6$). There was a more marked degradation of skill between training and testing time periods in the Lower Bay ($R^2 = 0.55$) and Patuxent River ($R^2 = 0.42$) (**Table 2**). Some point locations in the Bay were sampled repeatedly over the 30-year time series, and contained >500 observations. The skill of the models on unseen test data was similar whether the whole zone (e.g. Mid Bay) was aggregated together, or if these long-term station locations within each zone were assessed separately. For example, a station at -76.292°W , 38.319°N in the Mid Bay (CB5.1: **Fig. 6**) with 634 observations in the dataset had an out-of-model validation R^2 between observed and modeled surface salinity of 0.61: the same value for the Mid Bay as a whole (**Table 2**). Similarly, a station at -76.602°W , 38.425°N with 564 observations in the Patuxent River (LE1.1) had an R^2 between observed and modeled surface salinity of 0.40, whereas the value for the River as a whole was 0.42. The models were thus capturing seasonal and interannual variability at point locations in the Chesapeake Bay with acceptable skill, and not simply reflecting (for example) climatological salinity gradients within specific zones.

To further assess the ability of the model trees to capture not only climatological patterns, but anomalous years, spatial fields of both temperature and salinity were compared for a cool September (2011) vs. a warm September (2008), and a dry September (2010) vs. a wet September (2011) (**Fig. 7**). September was chosen as it was a generally well-sampled month across the time series, both spatially and temporally. We note that September 2011 was particularly wet due to the effects of Tropical Storm Lee (Cheng et al. 2013). Hindcasts from the model trees captured not only the higher surface temperatures during a warm year (**Fig. 7c,d**), and lower salinities during a wet year (**Fig. 7g,h**), but also the general spatial patterns of these phenomena. In particular, both observations and models highlighted the downriver and down-Bay movement of isohalines during times of high Susquehanna River flow (**Fig 7g,h**). The surface temperature predictions were somewhat more biased than those for surface salinity, with the western rivers showing a cool bias in 2011, and the Upper Bay showing a warm bias in 2008 (**Fig. 7a-d**). However, the overall spatial structure was reproduced reasonably well.

Projected atmospheric temperature and precipitation changes

Comparison of the four statistical downscaling methods used to link projected GCM-scale changes in air temperature and precipitation to the Chesapeake Bay showed that downscaled changes were similar in magnitude to those from the corresponding GCMs (2 – 5.5 °C, ~0 - 200 mm year⁻¹ rainfall) (**Fig. 8**). However, overall trends and separation amongst the 4 GCMs was less clear for precipitation, which showed much stronger interannual variability than temperature. In addition, the contribution of downscaling methods to variability in projections was much less than the contribution of inter-GCM variability, for the 4 GCMs examined here (**Fig. 8**). Choice of downscaling method, however, could exert a significant impact by the end of the 21st century for GCMs with the largest projected warming. The mean difference among the warmest and coolest downscaling methods for 2071 - 2100 was 0.7°C for the WW model, and 0.8°C for the WD model (**Fig. 8a**).

Surface temperature and salinity futures for the Chesapeake Bay

Annual mean projections of Bay-wide averages of surface temperature showed a clear warming trend similar to that of air temperature (**Fig. 9a**). Projections of surface salinity and Susquehanna River streamflow were more variable, however the two wetter models (WW and CW) showed little change to annual mean salinity or streamflow by the end of the century. In contrast, the two dry models (WD and CD) projected significantly decreased mean annual streamflow, and thus increasing surface salinity (**Fig.**

9b,c). When this analysis was repeated for each season (graph not shown), the CW model also showed a significant decrease in winter salinity, and increase in streamflow, by 2100 (linear regression, $p < 0.05$, Durbin-Watson test $p > 0.05$)

Projections from each GCM averaged across all downscaling methods showed warming of Chesapeake Bay surface waters during all seasons by the latter half of the last 30 years of the 21st century (**Fig. 10**). However, the extent of overall warming and its seasonal distribution varied among models. The WW model projected the strongest warming in estuarine water temperatures between 1970 - 1999 and 2071 - 2100: $> 5^{\circ}\text{C}$ in all seasons. Warming in this model was particularly strong in winter (January - March) and fall (October - November), when compared to the other GCMs. In contrast, the WD model showed the weakest warming in winter: 2.7°C between 1970 - 1999 and 2070 - 2100, vs $>4^{\circ}\text{C}$ in all other seasons. The CW model projected the weakest warming overall: from 2.3°C in winter to 3.0°C in fall by the end of the century. Both the WW and WD models projected mean summer (July - September) surface temperatures across the Bay of $>30^{\circ}\text{C}$ by 2071 - 2100, compared to a historical (1970 - 1999) mean of $25.5 - 26^{\circ}\text{C}$. In contrast, the CW model projected mean summer water temperatures of 28.0°C for 2071 - 2100, and the CD model 28.9°C . The strong winter warming in the WW model resulted in projections of mean winter surface temperatures of 11.5°C , much warmer than for 1970 - 1999, where the mean was 6.3°C . In contrast, the CW, WD and CD models projected winter surface temperatures of $8.7 - 9.0^{\circ}\text{C}$ by the end of the century (**Fig. 10**). Projections of winter warming and streamflow changes were more variable than those for summer, across all models.

Warming in the WW and WD models resulted in projected conditions well outside the range of historical variability during summer. For example, the maximum observed surface water temperature at the Thomas Point buoy (1985 - 2015) was 29.9°C , recorded on August 4th, 2006. By the end of the 21st century (2071-2100), this value was projected to be exceeded on 63% of days in July, in both the WW and WD models. For August, this value was exceeded on 100% of days in the WW model, and 87% of days in the WD model. In contrast, surface temperature at Thomas Point was projected to exceed 29.9°C on only 14% of days in August in the CW model by the end of the century, and 37.6% of days in the CD model. Thus summer conditions in the Chesapeake Bay were projected to be strongly novel in the WW and WD models, but less so in the CW and CD models.

Changes in streamflow varied strongly among models. While the wetter WW and CW models projected increases in late winter/early spring streamflow from the Susquehanna River and stable mean flows during other seasons, the WD model projected decreased streamflow across most seasons by the end of the century (**Fig. 10**). This was despite modest increases in precipitation, and was due to increased evapotranspiration associated with warming atmospheric conditions. The different models were in much closer agreement for summer, with all showing either stable or slightly decreasing streamflow. Surface salinity changes reflected differences in streamflow patterns across models. Changes were close to neutral for much of the spring through fall in all models. Projections diverged, however, in the winter and early spring, with more saline conditions prevailing for the drier WD model, and fresher conditions for the WW and CW models.

Overall, choice of statistical downscaling method contributed less to variability in projections than choice of GCM (**Fig. 8**). However, a comparison of the coolest (BCQM) and warmest (CFQM) methods for the WW model highlighted considerable discrepancies in projections at certain times of year (**Fig. 11**). The CFQM model was 1.4°C warmer than the BCQM model during the summer, and Susquehanna River streamflow was ~200 m³ s⁻¹ less during winter-early spring (January – April). Differences in salinity between the two downscaling methods were small, but were slightly lower for BCQM.

Projected increases in surface temperature were largely spatially coherent across the Chesapeake Bay, but some spatial structure was apparent. We focus on the summer period, when the Chesapeake Bay is most likely to experience warm conditions beyond previously recorded highs, which could stress the physiological limits of contemporary marine communities. Results from other seasons are shown in the Supplementary Material (**Fig. S1**). In each case, the two GCMs with the smallest and largest changes in temperature are shown. Reduced warming was associated with the southern end of the Chesapeake Bay, due to the moderating influence of continental shelf waters (**Fig. 12**). Maximum warming was associated with the Upper Bay, particularly in the WW model.

Stronger spatial structure was apparent in the projected changes in surface salinity. We focus on the winter period (January – March) where projected changes in streamflow were highest, as were contrasts between the models. Results from other seasons are shown in the Supplementary Material (**Fig. S2**). In each case, the two GCMs with the smallest and largest changes in salinity are shown. Changes in the WW model were strongest in the Upper Bay, and in mid-low reaches of some western rivers. These are

currently transition zones between oligohaline and mesohaline waters, or mesohaline and polyhaline waters, and salinity decreases in these areas represent mean downstream movement of isohalines. The weakest salinity change was in upstream portions of rivers, where conditions are currently tidal fresh to oligohaline, and so increasing streamflow cannot decrease salinity substantially (**Fig. 13**). In contrast, the decreased streamflow in the WD model lead to projected increases in salinity throughout the Bay. Similarly to the WW model, changes were strongest in the Upper Bay, and the mid reaches of some western rivers, representing a particularly large percentage change given current low salinities in these areas. Weakest changes in salinity were projected for tidal fresh to oligohaline reaches of rivers, which were expected to stay largely fresh despite reduced streamflow.

Discussion

Modeling framework: uncertainty and complexity

A key challenge for understanding potential climate change impacts on natural resources and environments is the development of projections at the appropriate spatial scale. GCMs generally have coarse spatial resolution, and inherent bias in the simulation of important processes (Wood et al. 2004; Xu and Yang 2012). Downscaling using either dynamical or statistical techniques usually improves projections, but each family of methods is subject to advantages and disadvantages. Dynamical downscaling has the advantage of explicitly and mechanistically representing physical processes controlling regional climate (Hellström et al., 2001). However, it is computationally expensive, which makes comparison of multiple GCMs and emission scenarios more difficult. In addition, the dynamically downscaled model will usually inherit the bias of the parent GCM, and addressing this issue is complex (Xu and Yang 2012). This is an important consideration for projections involving hydrological simulations, which are sensitive to bias in both the mean and spatial distribution of watershed properties (Wood et al., 2004).

Statistical downscaling has the disadvantage of relying on empirical relationships between coarse- and fine-scale processes, and assuming that these relationships will remain valid when projected into the future (i.e. stationarity: Schmith 2008; Michelangeli et al. 2009; Cannon 2010; Kallache et al. 2011; Gaitán and Cannon 2013; Dixon et al. 2016). However, advantages include inherent bias correction, and a much lower computational cost than dynamical downscaling. As a result, the statistical framework presented here allowed consideration of a range of projected surface temperature and salinity futures for the Chesapeake Bay under climate change. While relatively simple, our approach was able to

reproduce historical conditions well at monthly resolution, and facilitated the easy and rapid comparison of multiple GCMs and statistical downscaling methods. To establish the methodology, here we tested four GCMs spanning the range of plausible future temperature and precipitation projections from CMIP5, and one CO₂ concentration pathway, however the statistical framework could easily ingest outputs from other models and scenarios.

We build on previous studies of potential climate change impacts to the Chesapeake Bay by using statistical model trees to consider both surface temperature and salinity in two dimensions. Although the use of one-dimensional air temperature as a proxy for water temperature (e.g. Pilgrim et al. 1998; Hare & Able 2007; Tisseuil et al. 2012; Jacobs et al. 2015) is a reasonable strategy in shallow estuarine environments, the more complex approach used in this study confers several advantages. Firstly, nonlinearities in the air-temperature vs. water-temperature relationship could be accounted for, in a way that allowed for future extrapolation. Secondly, the influence of streamflow on surface water temperature could be included, although this effect was minor compared to that of air temperature in the Chesapeake Bay. This may not be the case in other estuaries more influenced by snowmelt, however. More importantly, projections of Chesapeake Bay salinity changes exhibited spatial contrasts within the Bay which have not been previously estimated, though changes in estuarine salinity have effects on marine resources comparable to those imposed by more often emphasized temperature shifts (Rome et al., 2005; Constantin de Magny et al., 2009; Jacobs et al., 2014, see additional discussion below).

While more complex than a linear air temperature proxy model, our approach was much less complicated than many other examples in the literature. Our water balance model was one of the simplest available, and did not include complex soil dynamics, flow routing, nutrient and sediment dynamics, or groundwater inflow (e.g. Hayhoe et al. 2007; Chen et al. 2014; Demaria et al. 2016; Lee et al. 2016), which may have led to an over-estimation of watershed evapotranspiration (Milly and Dunne 2011), and an inability to capture extreme flow events (e.g. hurricanes, floods). We also did not consider hydrodynamics of the estuary in three dimensions, which usually requires additional parameters, such as wind fields (Gibson and Najjar 2000; Valle-Levinson et al. 2001; Xu et al. 2012; Lee et al. 2013; Irby et al. 2016). Considering the effects of sea level rise on temperature and salinity in the Chesapeake Bay was also beyond the scope of this study, even though these may be substantial (Hong and Shen, 2012; Ross et al. 2015). This was largely because the impacts of sea level rise on estuarine conditions are

spatially and temporally complex (e.g. Hilton et al. 2008), and because projections of the future magnitude of sea level rise are extremely divergent and uncertain (Church et al. 2013; Grinsted et al. 2015). Our projections of surface temperature and salinity would also not represent potential changes throughout the entire water column, which are more relevant to organisms which do not live at the surface. Despite these simplifications, our approach was able to reproduce observations in the Chesapeake Bay at a monthly resolution with good accuracy, using only air temperature and watershed precipitation as inputs. However, different or modified models may be required for other watersheds with different characteristics (e.g., smaller, more arid, more influenced by snowmelt etc.). We also note the importance of continued development of dynamical approaches that allow extrapolation beyond surface properties and to biogeochemical impacts such as hypoxia (Bever et al. 2013; Brown et al. 2013; Testa et al. 2014; Feng et al. 2015). Applications of statistical and dynamical tools in concert may maximize the benefits of large ensembles that statistical approaches facilitate, along with the stronger mechanistic linkages enabled by dynamical approaches.

An advantage of the simplicity of our statistical framework was the ability to include projections from multiple GCMs, downscaled using multiple methods. We found that the choice of GCM contributed much more to overall uncertainty than choice of downscaling method. However, this was largely a product of the decision to include GCMs with widely diverging futures, but to use statistical downscaling methods with reasonably similar characteristics. Previous studies (e.g. Wood et al. 2004; Wilby and Harris 2006; Chen et al. 2011; Mandal et al. 2016) have found that the choice of statistical downscaling method can contribute considerable uncertainty to future projections in some systems. However, the suite of GCMs, hydrological models and statistical downscaling methods differed between each study. Clearly, the relative contribution of model and downscaling methods to projection uncertainty depends on both the method and model selected, and will be region specific (Johnson et al. 2012). A more exhaustive assessment of their relative influence on projection uncertainty will require consideration of a wider range of downscaling methods. Ideally, downscaling approaches would be explicitly considered within a full suite of uncertainty sources impacting projections (Hawkins and Sutton 2011; Cheung et al. 2016).

While in our study the choice of GCM was more influential than the choice of statistical downscaling method, the two warmer models (WW and WD) showed some divergence of projections from different downscaling methods in the later 21st century. Results diverged primarily at the tails of the distributions,

particularly where future conditions were outside ranges experienced in the historical period (e.g. summer temperatures: **Fig. S3**). This was primarily a result of the different bias-correction procedures used in each technique (see Methods). Many climate change impact studies include only one statistical downscaling method, without consideration of the effect of this choice on overall propagation of error or variance. Our results suggest that the choice of downscaling method can introduce considerable variability once conditions diverge from current observations, even among closely related methods. For example, projected mean annual air temperatures in the Susquehanna River watershed by the end of the 21st century in the WD model differed by 0.8°C between the four statistical downscaling methods. The more conservative EDQM method projected mean temperatures from 2071 – 2100 of 15.4°C, while the warmer CFQM method projected a mean of 16.2°C. The other methods (BCQM and CDFt) were intermediate between the two.

Potential impacts from projected temperature and salinity changes

Results from this study highlighted several potential changes to surface temperature and salinity in the Chesapeake Bay by the end of the century. All GCMs projected substantial warming under RCP8.5, however the magnitude of warming between the late 20th and late 21st centuries varied markedly. For example, the WW model projected an increase in summer surface temperatures of >5°C, while the CW model projected as little as ~2°C. The models also disagreed on which season would see the strongest warming. In all models, the magnitude of summer warming was more much less variable than that for winter. Changes to salinity were even less certain, particularly for winter and fall. While the warmer, drier WD model projected increases in salinity for both these seasons, the WW and CW models showed salinity decreases for winter. In contrast, all models projected weaker changes to salinity for summer.

These findings are broadly consistent with many of those from Najjar et al. (2010), who reviewed potential climate change impacts on the Chesapeake Bay using CMIP3 GCMs. Apart from one outlier (CCSR), a 7-member ensemble also projected a ~2 – 5.5°C increase in temperature over the Chesapeake Bay watershed area in this earlier study. However, similarly to the present study, precipitation projections were much more variable, ranging from a ~30% decrease to a ~20% increase within seasons. Variability in freshwater inflow to the Chesapeake Bay is largely driven by precipitation, rather than evapotranspiration (Najjar 1999; Gibson and Najjar 2000; Najjar et al. 2010). Uncertainty in precipitation projections from the GCMs is thus the main driver of uncertainty in projected streamflow, and surface salinity within the Bay. This uncertainty in projected future precipitation, and thus streamflow, is a

common thread in many other studies from the NE United States (Najjar et al. 2010; Johnson et al. 2012), and other locations (Schneider et al. 2013; Bosshard et al. 2014).

The key addition of our study relative to previous GCM syntheses (including Najjar et al. 2010) is inference of the implications of these large-scale changes on spatial temperature and salinity patterns in Chesapeake Bay. These are, to our knowledge, the first spatially resolved salinity and temperature projections for the Bay, and the first comparison of projection uncertainty at the scale of the estuary itself. For summer warming, results highlight a strong coherent warming signal with uncertainty bounds mimicking the range of projected changes in surface air temperature. Weaker warming was projected for the Lower Bay, potentially due to the moderating influence of nearby continental shelf waters. For salinity, results highlight regions likely subject to larger salinity changes. For the majority of projections showing the potential for altered streamflow during the winter and early spring, mesohaline regions are likely to experience the largest absolute changes in salinity, however oligohaline areas may experience the largest percentage changes.

The strong temperature and salinity gradients in the Chesapeake Bay result in distinct spatial distributions of resident organisms, driven by species-specific physical tolerances (Atwood et al. 2001; Cotton et al. 2003; Jung and Houde 2003; Kimmel et al. 2006). Climate-driven changes in temperature and salinity will therefore alter spatiotemporal habitat availability, and species may need to change their spatial distributions and migratory patterns if conditions begin to exceed tolerable ranges (Wood et al. 2002; Najjar et al. 2010). For example, some coldwater species such as winter flounder (*Pseudopleuronectes americanus*) are currently only present in the Chesapeake Bay during cooler months. Laboratory experiments on adults suggest a relatively cool temperature preference of 13 - 14°C, and they have been observed to stop feeding at temperature >23°C (Olla et al. 1969; Periera et al. 1999). Projections from this study suggest mean spring surface water temperatures of 21.5°C (CW) to >23°C (WW) by the end of the century under RCP8.5 (**Fig. 9**), compared to a recent historical mean of ~18°C. Under this scenario, if the rest of the water column warms at a comparable rate to surface waters, this species may spend less time in the Chesapeake Bay, or eventually be excluded altogether. This is particularly likely if future conditions follow projections suggested by the warmer WW and WD models. While projections from our study show some slight thermal refugia in the Lower Bay and lower James River, most of the Chesapeake Bay is projected to warm substantially. Similarly, juvenile Atlantic sturgeon (*Acipenser oxyrinchus*) are physiologically stressed by temperatures of >28°C in their first few

years of life in the Bay (Niklitschek and Secor 2005). While recent historical and present-day summer surface temperatures average ~25 - 26°C, these may increase to between 27 - 29°C (CW) and >30°C (WW) by the end of the century. The differences in projected temperature in the Chesapeake Bay among different climate models could thus encompass the difference between a moderate and potentially lethal change in conditions for some species.

Hales and Able (2001) showed that young-of-the-year of black sea bass (*Centropristis striata*) could not tolerate water temperatures of <2 - 3°C. This restriction is also shared by other species which occur in the Chesapeake Bay: summer flounder (*Paralichthys dentatus*), striped bass (*Morone saxatilis*), Atlantic croaker (*Micropogonias undulatus*), weakfish (*Cynoscion regalis*) and spot (*Leiostomus xanthurus*) also appear to have lethal lower temperature limits of ~2 - 3°C (Schwartz 1964; Malloy and Targett 1991; Atwood et al. 2001; Lankford and Targett 2001; Rome et al. 2005). In the case of blue crab (*Callinectes sapidus*), this temperature limit also interacts with salinity, with cold, low salinity conditions least favorable for survival (Rome et al. 2005; Hines et al. 2010). As a result of these tolerance limits, recruitment variability in some of these species (e.g. Atlantic croaker: Hare and Able 2007) has been directly linked to over-wintering mortality of juveniles in estuarine habitats, with warmer winters being more favorable. During the monitoring period of the Thomas Point buoy (1985-2016), mean monthly surface temperatures were <3°C in 48% of years for January, 54% of years for February, and 3% of years for December. By 2071-2100, projected mean monthly surface temperature at the Thomas Point location fell below 3°C on only one occasion: a January in the CW GCM. Our results thus suggest that under RCP8.5, even using the most conservative GCM considered, the over-wintering mortality restriction on recruitment for many fish species may be completely removed in the Chesapeake Bay.

Species that rely on environmental cues for spawning initiations and migration may also shift their phenology in response to changing temperature and flow characteristics. For example, temperature influences movement of striped bass within the Chesapeake Bay, and both temperature and flow regimes drive immigration and emigration of river herring (*Alosa aestivalis*, *A. pseudoharengus*) in and out of the Bay, across multiple life stages (Peer and Miller 2015; Tommasi et al. 2015). Adult spawning activity may be associated with specific conditions most favorable for larval survival: larval striped bass survive best at 15 - 20°C, and current spawning activity peaks from April-June, where surface temperatures currently average ~18°C across the Chesapeake Bay (Rutherford and Houde 1995; Secor and Houde 1995). These are projected to increase to between 21.4°C (CW) and 23.3°C (WW) by the end

of the century. To ensure spawning success, striped bass will therefore have to acclimate to warmer temperatures, or shift their spawning season to earlier in the year. Spawning activity in some species is also timed to maximize food availability for larvae, and to take advantage of seasonal blooms in primary productivity caused by temperature and flow patterns. If the timing of favorable temperatures shifts at a different rate to the timing of favorable feeding conditions, a mismatch may result, with implications for recruitment success (Wood et al. 2002).

Future temperature and salinity changes will also impact important benthic habitats, such as submerged aquatic vegetation (SAV). In the Chesapeake Bay, these primarily comprise seagrasses in higher salinity zones, and freshwater angiosperms in lower salinity locations (Dennison et al. 1993; Kemp et al. 2004). SAVs provide essential habitat for multiple life stages of vertebrate and invertebrate organisms in the Bay, as well as assimilating nutrients and reducing turbidity (Kemp et al. 2005). However, they are strongly sensitive to changes in salinity regimes and water quality within the estuary. Future changes to salinity in the Bay may shift the spatial distributions of different types of SAV, depending on their physiological tolerances. In addition, any decrease in water quality driven by changing streamflow regimes (e.g. Lee et al. 2016) may reduce light penetration, and lead to the loss of SAV beds.

Climate change is likely to influence the abundance and distribution of not just economically and ecologically beneficial species, but also nuisance and pathogenic organisms which lead to management issues. A good example of this is the Chesapeake Bay is the occurrence of *Vibrio* spp., which cause potentially severe illness in humans through foodborne and environmental exposure (Ralston et al. 2011). *Vibrios* are currently most abundant in warmer water temperatures in the Chesapeake Bay, and associated with species-specific salinity ranges (Kaneko and Colwell 1973, 1978; Constantin de Magny et al. 2009; Jacobs et al. 2014, 2015). Laboratory experiments show optimum temperatures for these species of 37 - 39°C: much warmer than for the vertebrates described above, and much warmer than currently observed water temperatures in the Chesapeake Bay (Kelly 1982; Miles et al. 1997; Sedas 2007). Projections from the statistical framework therefore suggest that *Vibrio* concentrations in the Bay are likely to increase. Jacobs et al. (2014) found that 96% of positive water samples for *V. vulnificus* were collected at temperatures warmer than 15°C. Surface temperature in the Chesapeake Bay is currently above this threshold between May and October, on average. By the end of the century, the CW model projected an extension of this time period to April through October, while the WW, WD and CD models projected favorable conditions for April through November. Changes to projected surface salinity varied

among models, but the increasing summer salinity in tributaries projected by the WD and CW models would also likely move *Vibrio* hotspots upstream.

In conclusion, we found that an empirical modeling framework using statistically downscaled air temperature and precipitation was able to reproduce monthly historical flow, surface temperature, and surface salinity characteristics in the Chesapeake Bay. While choice of GCM contributed a large amount of uncertainty to future projections, downscaled global climate models suggest a 2 - 5.5°C increase in surface water temperatures in the Chesapeake Bay by the end of the century in all seasons. Projections of streamflow were more uncertain, but may increase in the winter and spring, and decrease in the fall, with subsequent impacts on surface salinity. These changes have implications for biological organisms that currently use the Bay as feeding, spawning, or nursery habitat, particularly those that are currently approaching their upper thermal limits during summer. In contrast, limits to recruitment on several species currently imposed by cold winters may be largely removed. There were multiple uncertainties associated with our study, including the simplification or exclusion of important physical and biological processes. However, results presented here provide a simple starting point for investigation of climate change impacts on spatial characteristics of the Chesapeake Bay, and potentially other estuaries around the world.

Acknowledgements

The authors wish to acknowledge T. Miller (University of Maryland) and M. Fabrizio, R. Latour, D. Kaplan, C. Meynard and D. Gauthier (Virginia Institute of Marine Science) for the provision of observational data in the Chesapeake Bay. H. Townsend, T. Ihde, B. Kinlan, M. Monaco and J. Manderson contributed valuable advice and discussion on data and analyses. J. Lanzante advised on the implementation of quantile mapping methods. The manuscript was significantly improved by review from A. Muñoz, M. Lee and two anonymous reviewers. GCMs were assessed using output from the NOAA ESRL climate change portal (J. Scott and M. Alexander). CTD data were obtained from the Chesapeake Bay Program's water quality database, with help from M. Mallonee. Air temperature observations were obtained from the NOAA National centers for Environmental Information, and Thomas Point observations from the NOAA National Data Buoy Center. CPC U.S. Unified Precipitation data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>. Funding and support for this study was provided by the NOAA National Ocean Service (NOS) National Centers for Coastal Ocean Science (NCCOS), the NOAA National Marine Fisheries Service (NMFS) Office of Science &

Technology, the NOAA Integrated Ecosystem Assessment (IEA) Program, and the NOAA office of Oceanic and Atmospheric Research (OAR).

References

- Adam, J.C. and D.P. Lettenmaier. 2003. Adjustment of global gridded precipitation for systematic bias. *Journal of Geophysical Research: Atmospheres* 108: doi:10.1029/2002JD002499.
- Atwood, H.L., S.P. Young, J.R. Tomasso Jr and T.I. Smith. 2001. Salinity and temperature tolerances of black sea bass juveniles. *North American Journal of Aquaculture* 63: 285-288.
- Barbier, E.B., S.D. Hacker, C. Kennedy, E.W. Koch, A.C. Stier and B.R. Silliman. 2011. The value of estuarine and coastal ecosystem services. *Ecological Monographs* 81: 169-193.
- Beck, M.W., K.L. Heck Jr, K.W. Able, D.L. Childers, D.B. Eggleston, B.M. Gillanders, B. Halpern, C.G. Hays, K. Hoshino, T.J. Minello and R.J. Orth. 2001. The identification, conservation, and management of estuarine and marine nurseries for fish and invertebrates: a better understanding of the habitats that serve as nurseries for marine species and the factors that create site-specific variability in nursery quality will improve conservation and management of these areas. *Bioscience* 51: 633-641.
- Bell, R.J., J.A. Hare, J.P. Manderson and D.E. Richardson. 2014. Externally driven changes in the abundance of summer and winter flounder. *ICES Journal of Marine Science* 71: 2416-2428.
- Benestad, R.E., I. Hanssen-Bauer and D. Chen. 2008. Empirical-statistical downscaling. World Scientific Publishing Company Incorporated, New Jersey.
- Bever, A.J., M.A.M. Friedrichs, C.T. Friedrichs, M.E. Scully, L.W.J. Lanerolle. 2013. Combining observations and numerical model results to improve estimates of hypoxic volume within the Chesapeake Bay, USA. *Journal of Geophysical Research: Oceans* 118: 4924-4944.
- Boesch, D.F., R.B. Brinsfield and R.E. Magnien. 2001. Chesapeake Bay eutrophication. *Journal of Environmental Quality* 30: 303-320.
- Bosshard, T., S. Kotlarski, M. Zappa and C. Schär. 2014. Hydrological climate-impact projections for the Rhine River: GCM–RCM uncertainty and separate temperature and precipitation effects. *Journal of Hydrometeorology* 15: 697-713.
- Breitburg, D.L. 1990. Near-shore hypoxia in the Chesapeake Bay: Patterns and relationships among physical factors. *Estuarine, Coastal and Shelf Science* 30: 593-609.
- Brown, C.W., R.R. Hood, W. Long, J. Jacobs, D.L. Ramers, C. Wazniak, J.D. Wiggert, R. Wood and J. Xu. 2013. Ecological forecasting in Chesapeake Bay: Using a mechanistic-empirical modeling approach. *Journal of Marine Systems* 125: 113-125.

833 Brown, L.R., L.M. Komoroske, R.W. Wagner, T. Morgan-King, J.T. May, R.E. Connon and N.A. Fangué.
834 2016. Coupled downscaled climate models and ecophysiological metrics forecast habitat
835 compression for an endangered estuarine fish. *PLoS ONE* 11: e0146724.
836 doi:10.1371/journal.pone.0146724

837 Cannon, A.J. and P.H. Whitfield. 2002. Downscaling recent streamflow conditions in British Columbia,
838 Canada using ensemble neural network models. *Journal of Hydrology* 259: 136-151.

839 Cannon, A.J. 2010. A flexible nonlinear modelling framework for nonstationary generalized extreme
840 value analysis in hydroclimatology. *Hydrological Processes* 24: 673-685.

841 Cannon, A.J. 2011. Quantile regression neural networks: Implementation in R and application to
842 precipitation downscaling. *Computers and Geosciences* 37: 1277-1284.

843 Chen, M., W. Shi, P. Xie, V. Silva, V.E. Kousky, R.W. Higgins and J.E. Janowiak. 2008. Assessing objective
844 techniques for gauge-based analyses of global daily precipitation. *Journal of Geophysical Research:*
845 *Atmospheres* 113: doi:10.1029/2007JD009132

846 Chen, J., F.P. Brissette, A. Poulin and R. Leconte. 2011. Overall uncertainty study of the hydrological
847 impacts of climate change for a Canadian watershed. *Water Resources Research* 47:
848 doi:10.1029/2011WR010602.

849 Chen, J. F.P. Brissette, D. Chaumont, M. Braun. 2013. Performance and uncertainty evaluation of
850 empirical downscaling methods in quantifying the climate change impacts on hydrology over two
851 North American river basins. *Journal of Hydrology* 479: 200-214.

852 Chen, X., K. Alizad, D. Wang and S.C. Hagen. 2014. Climate change impact on runoff and sediment loads
853 to the Apalachicola River at seasonal and event scales. *Journal of Coastal Research* 68: 35-42.

854 Cheng, P., Li, M., Li, Y. 2013. Generation of an estuarine sediment plume by a tropical storm. *Journal of*
855 *Geophysical Research: Oceans* 118: 856-868.

856 Cheung, W.W., M.C. Jones, G. Reygondeau, C.A. Stock, V.W. Lam and T.L. Frölicher. 2016. Structural
857 uncertainty in projecting global fisheries catches under climate change. *Ecological Modelling* 325:
858 57-66.

859 Church, J.A., P.U. Clark, A. Cazenave, J.M. Gregory, S. Jevrejeva, A. Levermann, M.A. Merrifield, G.A.
860 Milne, R.S. Nerem, P.D. Nunn, A.J. Payne, W.T. Pfeffer, D. Stammer, A.S. Unnikrishnan. 2013: Sea
861 Level Change. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I
862 to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D.
863 Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, P.M. Midgley (eds.)].
864 Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Cloern, J.E., P.C. Abreu, J. Carstensen, L. Chauvaud, R. Elmgren, J. Grall, H. Greening, J.O.R. Johansson, M. Kahru, E.T. Sherwood and J. Xu. 2016. Human activities and climate variability drive fast-paced change across the world's estuarine–coastal ecosystems. *Global Change Biology* 22: 513-529.

Constantin De Magny, G., W. Long, C.W. Brown, R.R. Hood, A. Huq, R. Murtugudde and R.R. Colwell. 2009. Predicting the distribution of *Vibrio* spp. in the Chesapeake Bay: a *Vibrio cholerae* case study. *EcoHealth* 6: 378-389.

Costa, M.J., H.N. Cabral, P. Drake, A.N. Economou, C. Ferandez-Delgado, L. Gordo, J. Marchand and R. Thiel. 2002. Recruitment and production of commercial species in estuaries. In: Elliott, M. and K.L. Hemingway (eds) *Fishes in Estuaries*. Blackwell Science pp 54-123.

Cotton, C.F., R.L. Walker and T.C. Recicar. 2003. Effects of temperature and salinity on growth of juvenile black sea bass, with implications for aquaculture. *North American Journal of Aquaculture* 65: 330-338.

Demaria, E.M., R.N. Palmer and J.K. Roundy. 2016. Regional climate change projections of streamflow characteristics in the Northeast and Midwest US. *Journal of Hydrology: Regional Studies* 5: 309-323.

Dennison, W.C., R.J. Orth, K.A. Moore Jr., J.C. Stevenson, V. Carter, S. Kollar, P.W. Bergstrom, R.A. Batiuk. 2003. Assessing water quality with submerged aquatic vegetation. *Bioscience* 43: 86-94.

Diaz-Nieto, J. and R.L. Wilby. 2005. A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom. *Climatic Change* 69: 245-268.

Dixon, K.W., J.R. Lanzante, M.J. Nath, K. Hayhoe, A. Stoner, A. Radhakrishnan, V. Balaji and C.F. Gaitán. 2016. Evaluating the stationarity assumption in statistically downscaled climate projections: is past performance an indicator of future results? *Climatic Change* 135: 395-408.

Edgar, G. J., N.S. Barrett, D.J. Graddon and P.R. Last. 2000. The conservation significance of estuaries: a classification of Tasmanian estuaries using ecological, physical and demographic attributes as a case study. *Biological Conservation* 92: 383-397.

Elith, J., J.R. Leathwick and T. Hastie. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology* 77: 802-813.

Fan, Y. and H. van den Dool, H. 2008. A global monthly land surface air temperature analysis for 1948–present. *Journal of Geophysical Research: Atmospheres* 113: doi:10.1029/2007JD008470

Fasbender D. and T.B.M.J. Ouarda. 2010. Spatial Bayesian model for statistical downscaling of AOGCM to minimum and maximum daily temperatures. *Journal of Climate* 23: 5222-5242.

Feng, Y., M.A.M Friedrichs, J. Wilkin, H. Tian, Q. Yang, E.E. Hofmann, J.D. Wiggert, R.R. Hood. 2015. Chesapeake Bay nitrogen fluxes derived from a land-estuarine ocean biogeochemical modeling

897 system: Model description, evaluation, and nitrogen budgets. *Journal of Geophysical Research:*
898 *Biogeosciences* 120: 1666-1695.

899 Feyrer, F., J.E. Cloern, L.R. Brown, M.A. Fish, K.A. Hieb and R.D. Baxter. 2015. Estuarine fish communities
900 respond to climate variability over both river and ocean basins. *Global Change Biology* 21: 3608-
901 3619.

902 Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V.
903 Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason and M. Rummukainen.
904 2013. Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis.
905 Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on
906 Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y.
907 Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and
908 New York, NY, USA.

909 Gaitán, C.F., A.J. Cannon. 2013. Validation of historical and future statistically downscaled pseudo-
910 observed surface wind speeds in terms of annual climate indices and daily variability. *Renewable*
911 *Energy* 51: 489-496.

912 Gaitán, C.F., W.W. Hsieh, A.J. Cannon and P. Gachon. 2014. Evaluation of linear and non-linear
913 downscaling methods in terms of daily variability and climate indices: Surface temperature in
914 southern Ontario and Quebec, Canada. *Atmosphere-Ocean* 52: 211-221.

915 Gaitán, C.F. 2016. Effects of variance adjustment techniques and time-invariant transfer functions on
916 heat wave duration indices and other metrics derived from downscaled time-series. Study case:
917 Montreal, Canada. *Natural Hazards* 83: 1661-1681.

918 Gibson, J.R. and R.G. Najjar. 2000. The response of Chesapeake Bay salinity to climate-induced changes
919 in streamflow. *Limnology and Oceanography* 45: 1764-1772.

920 Gillanders, B.M., T.S. Elsdon, I.A. Halliday, G.P. Jenkins, J.B. Robins and F.J. Valesini. 2011. Potential
921 effects of climate change on Australian estuaries and fish utilising estuaries: a review. *Marine and*
922 *Freshwater Research* 62: 1115-1131.

923 Gleick, P.H. and D.B. Adams. 2000. Water: The potential consequences of climate variability and change
924 of the water resources of United States. The Report of the Water Sector Assessment Team of the
925 National Assessment of the Potential Consequences of Climate Variability and Change for the U.S.
926 Global Change Research Program. 160p

927 Glibert, P.M., R. Magnien, M.W. Lomas, J. Alexander, C. Tan, E. Haramoto, M. Trice and T.M. Kana. 2001.
 928 Harmful algal blooms in the Chesapeake and coastal bays of Maryland, USA: Comparison of 1997,
 929 1998, and 1999 events. *Estuaries* 24: 875-883.

930 Grinsted, A., S. Jevrejeva, R. E. Riva, D. Dahl-Jensen. 2015. Sea level rise projections for northern Europe
 931 under RCP8. 5. *Climate Research* 64: 15-23.

932 Guo, X. and A. Valle-Levinson. 2007. Tidal effects on estuarine circulation and outflow plume in the
 933 Chesapeake Bay. *Continental Shelf Research* 27: 20-42.

934 Hagy, J.D., W.R. Boynton, C.W. Keefe and K.V. Wood. 2004. Hypoxia in Chesapeake Bay, 1950–2001:
 935 long-term change in relation to nutrient loading and river flow. *Estuaries* 27: 634-658.

936 Hagy, J.D., W.R. Boynton and D.A. Jasinski. 2005. Modelling phytoplankton deposition to Chesapeake
 937 Bay sediments during winter–spring: interannual variability in relation to river flow. *Estuarine,
 938 Coastal and Shelf Science* 62: 25-40.

939 Hales Jr, L. and K. Able. 2001. Winter mortality, growth, and behavior of young-of-the-year of four
 940 coastal fishes in New Jersey (USA) waters. *Marine Biology* 139: 45-54.

941 Hare, J.A. and K.W. Able. 2007. Mechanistic links between climate and fisheries along the east coast of
 942 the United States: explaining population outbursts of Atlantic croaker (*Micropogonias undulatus*).
 943 *Fisheries Oceanography* 16: 31-45.

944 Hare, J.A., W.E. Morrison, M.W. Nelson, M.M. Stachura, E.J. Teeters, R.B. Griffis, M.A. Alexander, J.D.
 945 Scott, L. Alade, R.J. Bell, A.S. Chute et al. 2016. A vulnerability assessment of fish and invertebrates
 946 to climate change on the Northeast US Continental Shelf. *PLoS ONE* 11: e0146756.
 947 doi:10.1371/journal.pone.0146756

948 Hawkins, E. and R. Sutton. 2011. The potential to narrow uncertainty in projections of regional
 949 precipitation change. *Climate Dynamics* 37: 407-418.

950 Hayhoe, K., C.P. Wake, T.G. Huntington, L. Luo, M.D. Schwartz, J. Sheffield, E. Wood, B. Anderson, J.
 951 Bradbury, A. DeGaetano and T.J. Troy. 2007. Past and future changes in climate and hydrological
 952 indicators in the US Northeast. *Climate Dynamics* 28: 381-407.

953 Hellström, C., D. Chen, C. Achberger and J. Räisänen. 2001. Comparison of climate change scenarios for
 954 Sweden based on statistical and dynamical downscaling of monthly precipitation. *Climate Research*
 955 19: 45-55.

956 Hessami, M., P. Gachon, T.B. Ouarda and A. St-Hilaire. 2008. Automated regression-based statistical
 957 downscaling tool. *Environmental Modelling and Software* 23: 813-834.

958 Hilton, T.W., R.G. Najjar, L. Zhong, M. Li. 2008. Is there a signal of sea-level rise in Chesapeake Bay
 959 salinity? *Journal of Geophysical Research* 113: C09002, doi:10.1029/2007JC004247.

960 Hines, A.H., E.G. Johnson, M.Z. Darnell, D. Rittschof, T.J. Miller, L.J. Bauer, P. Rodgers and R. Aguilar.
 961 2010. Predicting effects of climate change on blue crabs in Chesapeake Bay. In: G.H. Kruse, G.L.
 962 Eckert, R.J. Foy, R.N. Lipcius, B. Sainte-Marie, D.L. Stram, and D. Woodby (eds.), *Biology and*
 963 *Management of Exploited Crab Populations under Climate Change*. Alaska Sea Grant, University of
 964 Alaska Fairbanks. doi:10.4027/bmecpcc.2010.22

965 Ho, C.K., D.B. Stephenson, M. Collins, C.A. Ferro and S.J. Brown. 2012. Calibration strategies: a source of
 966 additional uncertainty in climate change projections. *Bulletin of the American Meteorological Society*
 967 93: 21-26.

968 Hong, B., J. Shen. 2012. Responses of estuarine salinity and transport processes to potential future sea-
 969 level rise in the Chesapeake Bay. *Estuarine, Coastal and Shelf Science* 104: 33-45.

970 Howarth, R.W., A. Sharpley and D. Walker. 2002. Sources of nutrient pollution to coastal waters in the
 971 United States: Implications for achieving coastal water quality goals. *Estuaries* 25: 656-676.

972 Irby, I.D., M.A.M. Friedrichs, C.T. Friedrichs, A.J. Bever, R.R. Hood, L.W.J. Lanerolle, M. Li, Linker L., M.E.
 973 Scully, K. Sellner, J. Shen, J. Testa, H. Wang, P. Wang, M. Xia. 2016. Challenges associated with
 974 modeling low-oxygen waters in Chesapeake Bay: a multiple model comparison. *Biogeosciences* 13:
 975 2011-2028.

976 Jacobs, J.M., M. Rhodes, C.W. Brown, R.R. Hood, A. Leight, W. Long and R. Wood. 2014. Modeling and
 977 forecasting the distribution of *Vibrio vulnificus* in Chesapeake Bay. *Journal of Applied Microbiology*
 978 117: 1312-1327.

979 Jacobs, J., S.K. Moore, K.E. Kunkel and L. Sun. 2015. A framework for examining climate-driven changes
 980 to the seasonality and geographical range of coastal pathogens and harmful algae. *Climate Risk*
 981 *Management* 8: 16-27.

982 Jelks, H.L., S.J. Walsh, N.M. Burkhead, S. Contreras-Balderas, E. Diaz-Pardo, D.A. Hendrickson, J. Lyons,
 983 N.E. Mandrak, F. McCormick, J.S. Nelson and S.P. Platania. 2008. Conservation status of imperiled
 984 North American freshwater and diadromous fishes. *Fisheries* 33: 372-407.

985 Johnson, T.E. and C.P. Weaver 2009. A framework for assessing climate change impacts on water and
 986 watershed systems. *Environmental Management* 43: 118-134.

987 Johnson, T.E., J.B. Butcher, A. Parker and C.P. Weaver (2012) Investigating the sensitivity of U.S.
 988 streamflow and water quality to climate change: U.S. EPA Global Change Research Program's 20
 989 Watersheds project. *Journal of Water Resources Planning and Management* 138: 453-464.

990 Jung, S. and E.D. Houde. 2003. Spatial and temporal variabilities of pelagic fish community structure and
 991 distribution in Chesapeake Bay, USA. *Estuarine, Coastal and Shelf Science* 58: 335-351.
 992 Kallache, M., M. Vrac, P. Naveau and P.A. Michelangeli. 2011. Nonstationary probabilistic downscaling of
 993 extreme precipitation. *Journal of Geophysical Research: Atmospheres* 116: doi:
 994 10.1029/2010JD014892.
 995 Kaneko, T. and R.R. Colwell. 1973. Ecology of *Vibrio parahaemolyticus* in Chesapeake Bay. *Journal of*
 996 *Bacteriology* 113: 24-32.
 997 Kaneko, T. and R.R. Colwell. 1978. The annual cycle of *Vibrio parahaemolyticus* in Chesapeake Bay.
 998 *Microbial Ecology* 4: 135-155.
 999 Kelly, M.T. 1982. Effect of temperature and salinity on *Vibrio* (Benecke) *vulnificus* occurrence in a Gulf
 1000 Coast environment. *Applied and Environmental Microbiology* 44: 820-824.
 1001 Kemp, W.M., R. Batiuk, R. Bartleson, P. Bergstrom, V. Carter, C.L. Gallegos, W. Hunley, L. Karrh, E.W.
 1002 Koch, J.M. Landwehr, K.A. Moore, L. Murray, M. Naylor, N.B. Rybicki, J.C. Stevenson, D.J. Wilcox.
 1003 2004. Habitat requirements for submerged aquatic vegetation in Chesapeake Bay: Water quality,
 1004 light regime and physical-chemical factors. *Estuaries* 27: 363-377.
 1005 Kemp, W.M., W.R. Boynton, J.E. Adolf, D.F. Boesch, W.C. Boicourt, G. Brush, J.C. Cornwell, T.R. Fisher,
 1006 P.M. Glibert, J.D. Hagy and L.W. Harding. 2005. Eutrophication of Chesapeake Bay: historical trends
 1007 and ecological interactions. *Marine Ecology Progress Series* 303: 1-29.
 1008 Kimmel, D.G., W.D. Miller and M.R. Roman. 2006. Regional scale climate forcing of mesozooplankton
 1009 dynamics in Chesapeake Bay. *Estuaries and Coasts* 29: 375-387.
 1010 Kimmel, D.G and M.R. Roman. 2004. Long-term trends in mesozooplankton abundance in Chesapeake
 1011 Bay, USA: influence of freshwater input. *Marine Ecology Progress Series* 267: 71-83.
 1012 Kuhn, M., S. Weston, C. Keefer and N. Coulter. 2015. Cubist: Rule- and Instance-Based Regression
 1013 Modeling. R package version 0.0.18. <http://CRAN.R-project.org/package=Cubist>
 1014 Langland, M.J., S.W. Phillips, J.P. Raffensperger and D.L. Moyer. 2004. Changes in streamflow and water
 1015 quality in selected nontidal sites in the Chesapeake Bay Basin, 1985–2003. US Geological Survey
 1016 Scientific Investigations Report 2004-5259, 50p.
 1017 Lankford Jr, T.E. and T.E. Targett. 2001. Low-temperature tolerance of age-0 Atlantic croakers:
 1018 recruitment implications for US mid-Atlantic estuaries. *Transactions of the American Fisheries*
 1019 *Society* 130: 236-249.
 1020 Larsen, L.W. and E.L. Peck. 1974. Accuracy of precipitation measurements for hydrologic modeling.
 1021 *Water Resources Research* 10: 857-863.

1022 Lee, Y. J., W.R. Boynton, M. Li and Y. Li. 2013. Role of late winter–spring wind influencing summer
 1023 hypoxia in Chesapeake Bay. *Estuaries and Coasts* 36: 683-696.

1024 Lee, M., E. Shevliakova, S. Malyshev, P.C.D. Milly and P.R. Jaffé. 2016. Climate variability and extremes,
 1025 interacting with nitrogen storage, amplify eutrophication risk. *Geophysical Research Letters* 43:
 1026 doi:10.1002/2016GL069254.

1027 Letcher, B.H., D.J. Hocking, K. O’Neil, A.R. Whiteley, K.H. Nislow and M.J. O’Donnell. 2016. A hierarchical
 1028 model of daily stream temperature using air-water temperature synchronization, autocorrelation,
 1029 and time lags. *PeerJ* 4: e1727 <https://doi.org/10.7717/peerj.1727>.

1030 Li, H., J. Sheffield and E.F. Wood. 2010. Bias correction of monthly precipitation and temperature fields
 1031 from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching.
 1032 *Journal of Geophysical Research: Atmospheres* 115: doi: 10.1029/2009JD012882.

1033 Limburg, K.E. and J.R. Waldman. 2009. Dramatic declines in North Atlantic diadromous fishes. *BioScience*
 1034 59: 955-965.

1035 Lynch, P.D., J.A. Nye, J.A. Hare, C.A. Stock, M.A. Alexander, J.D. Scott, K.L. Curti and K. Drew. 2014.
 1036 Projected ocean warming creates a conservation challenge for river herring populations. *ICES*
 1037 *Journal of Marine Science* 72: 374-387.

1038 Malloy, K.D. and T.E. Targett. 1991. Feeding, growth and survival of juvenile summer flounder
 1039 *Paralichthys dentatus*: experimental analysis of the effects of temperature and salinity. *Marine*
 1040 *Ecology Progress Series* 72: 213-223.

1041 Mandal, S., P.A. Breach and S.P. Simonovic. 2016. Uncertainty in Precipitation Projection under Changing
 1042 Climate Conditions: A Regional Case Study. *American Journal of Climate Change* 5: 116-132.

1043 Martinez, M.L., A. Intralawan, G. Vazquez, O. Perez-Maqueo, P. Sutton and R. Landgrave. 2007 The
 1044 coasts of our world: Ecological, economic and social importance. *Ecological Economics* 63: 254-272.

1045 Maurer, E.P. and P.B. Duffy. 2005. Uncertainty in projections of streamflow changes due to climate
 1046 change in California. *Geophysical Research Letters* 32: doi:10.1029/2004GL021462.

1047 McCabe, G.J. and S.L. Markstrom. 2007. A monthly water-balance model driven by a graphical user
 1048 interface: U.S. Geological Survey Open-File report 2007-1088, 6 p.

1049 Michelangeli, P.A., M. Vrac and H. Loukos, H. 2009. Probabilistic downscaling approaches: Application to
 1050 wind cumulative distribution functions. *Geophysical Research Letters* 36: doi:
 1051 10.1029/2009GL038401.

1052 Miles, D.W., T. Ross, J. Olley and T.A. McMeekin. 1997. Development and evaluation of a predictive
 1053 model for the effect of temperature and water activity on the growth rate of *Vibrio*
 1054 *parahaemolyticus*. *International Journal of Food Microbiology* 38: 133-142.

1055 Milly, P.C.D and K.A. Dunne. 2011. On the hydrologic adjustment of climate-model projections: The
 1056 potential pitfall of evapotranspiration. *Earth Interactions* 15: doi:10.1175/2010EI363.1

1057 Milly, P.C., K.A. Dunne and A.V. Vecchia. 2005. Global pattern of trends in streamflow and water
 1058 availability in a changing climate. *Nature* 438: 347-350.

1059 Mohseni, O., H.G. Stefan and J.G. Eaton. 2003. Global warming and potential changes in fish habitat in
 1060 US streams. *Climatic Change* 59: 389-409.

1061 Najjar, R.G. 1999. The water balance of the Susquehanna River Basin and its response to climate change.
 1062 *Journal of Hydrology* 219: 7-19.

1063 Najjar, R.G., C.R. Pyke, M.B. Adams, D. Breitburg, C. Hershner, M. Kemp, R. Howarth, M.R. Mulholland,
 1064 M. Paolisso, D. Secor and K. Sellner. 2010. Potential climate-change impacts on the Chesapeake Bay.
 1065 *Estuarine, Coastal and Shelf Science* 86: 1-20.

1066 Niklitschek, E.J. and D.H. Secor. 2005. Modeling spatial and temporal variation of suitable nursery
 1067 habitats for Atlantic sturgeon in the Chesapeake Bay. *Estuarine, Coastal and Shelf Science* 64: 135-
 1068 148.

1069 Olla, B.L., R. Wicklund and S. Wilk. 1969. Behavior of winter flounder in a natural habitat. *Transactions of*
 1070 *the American Fisheries Society* 98: 717-720.

1071 Overland, J.E., M. Wang, N.A. Bond, J.E. Walsh, V.M. Kattsov and W.L. Chapman. 2011. Considerations in
 1072 the selection of global climate models for regional climate projections: the Arctic as a case study.
 1073 *Journal of Climate* 24: 1583-1597.

1074 Paerl, H.W. 2006. Assessing and managing nutrient-enhanced eutrophication in estuarine and coastal
 1075 waters: Interactive effects of human and climatic perturbations. *Ecological Engineering* 26: 40-54.

1076 Paerl, H.W. and T.G. Otten. 2013. Harmful cyanobacterial blooms: causes, consequences, and controls.
 1077 *Microbial Ecology* 65: 995-1010.

1078 Payne, M.R., M. Barange, W.W. Cheung, B.R. MacKenzie, H.P. Batchelder, X. Cormon, T.D. Eddy, J.A.
 1079 Fernandez, A.B. Hollowed, M.C. Jones and J.S. Link. 2016. Uncertainties in projecting climate-change
 1080 impacts in marine ecosystems. *ICES Journal of Marine Science* 73: 1272-1282.

1081 Peer, A.C. and T.J. Miller. 2015. Climate change, migration phenology, and fisheries management
 1082 interact with unanticipated consequences. *North American Journal of Fisheries Management* 34: 94-
 1083 110.

1084 Periera, J.J., R. Goldberg, J.J. Ziskowski, P.L. Berrien, W.W. Morse and D.L. Johnson. 1999. Winter
 1085 flounder, *Pseudopleuronectes americanus*, life history and habitat characteristics. NOAA Technical
 1086 Memorandum NMFS-NE-138, 48p.

1087 Pihl, L., A. Cattrijse, I. Codling, S. Mathieson, D.S. McLusky and C. Roberts. 2002. Habitat use by fishes in
 1088 estuaries and other brackish areas. In: Elliott, M. and K.L. Hemingway (eds) *Fishes in Estuaries*.
 1089 Blackwell Science pp 10-53.

1090 Pilgrim, J.M., X. Fang and H.G. Stefan (1998). Stream temperature correlations with air temperatures in
 1091 Minnesota: Implications for climate warming. *Journal of the American Water Resources Association*
 1092 34: 1109-1121.

1093 Preston, B.L. 2004. Observed winter warming of the Chesapeake Bay estuary (1949–2002): Implications
 1094 for ecosystem management. *Environmental Management* 34: 125-139.

1095 Quinlan, J.R. 1992. Learning with continuous classes. *5th Australian Joint Conference on Artificial*
 1096 *Intelligence* 92: 343-348.

1097 Quinlan, J.R. 1993. Combining instance-based and model-based learning. *Proceedings of the Tenth*
 1098 *International Conference on Machine Learning* pp. 236-243.

1099 R Core Team (2015). R: A language and environment for statistical computing. R Foundation for
 1100 Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

1101 Ralston, E.P., H. Kite-Powell and A. Beet. 2011. An estimate of the cost of acute health effects from food-
 1102 and water-borne marine pathogens and toxins in the USA. *Journal of Water and Health* 9: 680-694.

1103 Reay, W.G. and K.A. Moore. 2009. Introduction to the Chesapeake Bay National Estuarine Research
 1104 Reserve in Virginia. *Journal of Coastal Research* 57: 1-9.

1105 Richards, R.A. and P.J. Rago. 1999. A case history of effective fishery management: Chesapeake Bay
 1106 striped bass. *North American Journal of Fisheries Management* 19: 356-375.

1107 Rome, M.S., A.C. Young-Williams, G.R. Davis and A.H. Hines. 2005. Linking temperature and salinity
 1108 tolerance to winter mortality of Chesapeake Bay blue crabs (*Callinectes sapidus*). *Journal of*
 1109 *Experimental Marine Biology and Ecology* 319: 129-145.

1110 Ross, A.C., R.G. Najjar, M. Li, M.E. Mann, S.E. Ford, B. Katz. 2015. Sea-level rise and other influences on
 1111 decadal-scale salinity variability in a coastal plain estuary. *Estuarine, Coastal and Shelf Science* 157:
 1112 79-92.

1113 Rutherford, E.S. and E.D. Houde. 1995. The influence of temperature on cohort-specific growth, survival,
 1114 and recruitment of striped bass, *Morone saxatilis*, larvae in Chesapeake Bay. *Fishery Bulletin* 93: 315-
 1115 332.

1116 Scavia, D., J.C. Field, D.F. Boesch, R.W. Buddemeier, V. Burkett, D.R. Cayan, M. Fogarty, M.A. Harwell,
 1117 R.W. Howarth, C. Mason and D.J. Reed. 2002. Climate change impacts on US coastal and marine
 1118 ecosystems. *Estuaries* 25: 149-164.

1119 Schneider, C., C.L.R. Laize, M.C. Acreman and M. Flörke. 2013. How will climate change modify river flow
 1120 regimes in Europe? *Hydrology and Earth System Sciences* 17: 325-339.

1121 Schubel, J.R. and D.W. Pritchard. 1986. Responses of upper Chesapeake Bay to variations in discharge of
 1122 the Susquehanna River. *Estuaries* 9: 236-249.

1123 Schulte, J.A., R.G. Najjar and M. Li. 2016. The influence of climate modes on streamflow in the Mid-
 1124 Atlantic region of the United States. *Journal of Hydrology: Regional Studies* 5: 80-99.

1125 Schwartz, F.J. 1964. Effects of winter water conditions on fifteen species of captive marine fishes.
 1126 *American Midland Naturalist* 71: 434-444.

1127 Scott, J. D., M.A. Alexander, D.R. Murray, D. Swales and J. Eischeid. 2016. The Climate Change Web
 1128 Portal: A System to Access and Display Climate and Earth System Model Output from the CMIP5
 1129 Archive. *Bulletin of the American Meteorological Society* 97: 523-530.

1130 Secor, D.H. and E.D. Houde. 1995. Temperature effects on the timing of striped bass egg production,
 1131 larval viability, and recruitment potential in the Patuxent River (Chesapeake Bay). *Estuaries* 18: 527-
 1132 544.

1133 Sedas, V.T.P. 2007. Influence of environmental factors on the presence of *Vibrio cholerae* in the marine
 1134 environment: a climate link. *The Journal of Infection in Developing Countries* 1: 224-241.

1135 Seliger, H.H. and J.A. Boggs. 1988. Long term pattern of anoxia in the Chesapeake Bay. In M. Lynch and
 1136 E.C. Krome (eds.) *Understanding the Estuary: Advances in Chesapeake Bay Research*. Publication
 1137 Number 129 p. 570-583 Chesapeake Research Consortium, Solomons, Maryland.

1138 Sharov, A.F., J.H. Vølstad, G.R. Davis, B.K. Davis, R.N. Lipcius and M.M. Montane. 2003. Abundance and
 1139 exploitation rate of the blue crab (*Callinectes sapidus*) in Chesapeake Bay. *Bulletin of Marine Science*
 1140 72: 543-565.

1141 Sheffield, J., A.P. Barrett, B. Colle, D. Nelun Fernando, R. Fu, K.L. Geil, Q. Hu, J. Kinter, S. Kumar, B.
 1142 Langenbrunner and K. Lombardo. 2013. North American climate in CMIP5 experiments. Part I:
 1143 evaluation of historical simulations of continental and regional climatology. *Journal of Climate* 26:
 1144 9209-9245.

1145 Shen, J. and H.V. Wang. 2007. Determining the age of water and long-term transport timescale of the
 1146 Chesapeake Bay. *Estuarine, Coastal and Shelf Science* 74: 585-598.

1147 Schmith, T. 2008. Stationarity of regression relationships: Application to empirical downscaling. *Journal*
1148 *of Climate* 21: 4529-4537.

1149 Solomantine, D.P. and K.N. Dulal. 2003. Model trees as an alternative to neural networks in rainfall –
1150 runoff modelling. *Hydrological Sciences Journal* 48: 399-411.

1151 Sprague, L.A., M.J. Langland, S.E. Yochum, R.E. Edwards, J.D. Blomquist, S.W. Phillips, G.W. Shenk and
1152 S.D. Preston. 2000. Factors affecting nutrient trends in major rivers of the Chesapeake Bay
1153 watershed (No. 2000-4218). US Geological Survey Water Resources Investigation Report 2000-4218,
1154 109p.

1155 SRBC (2013) State of the Susquehanna: 2013 Report. Susquehanna River Basin Commission, 20pp

1156 Stock, C.A., M.A. Alexander, N.A. Bond, K.M. Brander, W.W. Cheung, E.N. Curchitser, T.L. Delworth, J.P.
1157 Dunne, S.M. Griffies, M.A. Haltuch and J.A. Hare. 2011. On the use of IPCC-class models to assess the
1158 impact of climate on living marine resources. *Progress in Oceanography* 88: 1-27.

1159 Tango, P.J., R. Magnien, W. Butler, C. Luckett, M. Luckenbach, R. Lacouture and C. Poukish. 2005.
1160 Impacts and potential effects due to *Prorocentrum minimum* blooms in Chesapeake Bay. *Harmful*
1161 *Algae* 4: 525-531.

1162 Testa, J.M., Y. Li, Y.J. Lee, M. Li, D.C. Brady, D.M. Di Toro, W.M. Kemp, J.J. Fitzpatrick. Quantifying the
1163 effects of nutrient loading on dissolved O₂ cycling and hypoxia in Chesapeake Bay using a coupled
1164 hydrodynamic-biogeochemical model. *Journal of Marine Systems* 139: 139-158.

1165 Thompson, J.R., A. Crawley and D.G. Kingston. 2015. GCM-related uncertainty for river flows and
1166 inundation under climate change: the Inner Niger Delta. *Hydrological Sciences Journal*
1167 doi:10.1080/02626667.2015.1117173

1168 Tisseuil, C., M. Vrac, G. Grenouillet, A.J. Wade, M. Gevrey, T. Oberdorff, J-B. Grodwohl and S. Lek. 2012.
1169 Strengthening the link between climate, hydrological and species distribution modeling to assess the
1170 impacts of climate change on freshwater biodiversity. *Science of the Total Environment* 424: 193-
1171 201.

1172 Tommasi, D., J. Nye, C. Stock, J.A. Hare, M. Alexander and K. Drew. 2015. Effect of environmental
1173 conditions on juvenile recruitment of alewife (*Alosa pseudoharengus*) and blueback herring (*Alosa*
1174 *aestivalis*) in fresh water: a coastwide perspective. *Canadian Journal of Fisheries and Aquatic*
1175 *Sciences* 72: 1037-1047.

1176 Valle-Levinson, A., K.C. Wong and K.T. Bosley. 2001. Observations of the wind-induced exchange at the
1177 entrance to Chesapeake Bay. *Journal of Marine Research* 59: 391-416.

1178 Vicuna, S. and J.A. Dracup. 2007. The evolution of climate change impact studies on hydrology and water
 1179 resources in California. *Climatic Change* 82: 327-350.
 1180 Viganò, G., G. Confortola, R. Fornaroli, R. Cabrini, S. Canobbio, V. Mezzanotte and D. Bocchiola. 2015.
 1181 Effects of Future Climate Change on a River Habitat in an Italian Alpine Catchment. *Journal of*
 1182 *Hydrologic Engineering* 21: 10.1061/(ASCE)HE.1943-5584.0001293, 04015063.
 1183 Vörösmarty, C.J., C.A. Federer and A.L. Schloss. 1998. Potential evaporation functions compared on US
 1184 watersheds: Possible implications for global-scale water balance and terrestrial ecosystem modeling.
 1185 *Journal of Hydrology* 207: 147-169.
 1186 Vrac, M. and P.A. Michelangeli. 2009. Package 'CDFt'. R package version 1.0.1 [https://cran.r-](https://cran.r-project.org/web/packages/CDFt/CDFt.pdf)
 1187 [project.org/web/packages/CDFt/CDFt.pdf](https://cran.r-project.org/web/packages/CDFt/CDFt.pdf)
 1188 Wagner, R.W., M. Stacey, L.R. Brown and M. Dettinger. 2011. Statistical models of temperature in the
 1189 Sacramento–San Joaquin Delta under climate-change scenarios and ecological implications.
 1190 *Estuaries and Coasts* 34: 544-556.
 1191 Wilber, D.H., D.G. Clarke, C.M. Alcoba and J. Gallo. 2016. Windowpane flounder (*Scophthalmus aquosus*)
 1192 and winter flounder (*Pseudopleuronectes americanus*) responses to cold temperature extremes in a
 1193 Northwest Atlantic estuary. *Journal of Sea Research* 107: 23-30.
 1194 Wilby, R.L. and T.M.L. Wigley. 1997. Downscaling general circulation model output: a review of methods
 1195 and limitations. *Progress in Physical Geology* 21: 530-548.
 1196 Wilby, R.L., C.W. Dawson and E.M. Barrow. 2002. SDSM—a decision support tool for the assessment of
 1197 regional climate change impacts. *Environmental Modelling and Software* 17: 145-157.
 1198 Wilby, R.L., O.J. Tomlinson and C.W. Dawson. 2003. Multi-site simulation of precipitation by conditional
 1199 resampling. *Climate Research* 23: 183-194.
 1200 Wilby, R. L. and I. Harris. 2006. A framework for assessing uncertainties in climate change impacts:
 1201 Low-flow scenarios for the River Thames, UK. *Water Resources Research* 42:
 1202 doi:10.1029/2005WR004065.
 1203 Wood, R.J., D.F. Boesch and V.S. Kennedy. 2002 Future consequences of climate change for the
 1204 Chesapeake Bay ecosystem and its fisheries. *American Fisheries Society Symposium* 32: 171-184.
 1205 Wood, A.W., L.R. Leung, V. Sridhar and D.P. Lettenmaier. 2004. Hydrologic implications of dynamical and
 1206 statistical approaches to downscaling climate model outputs. *Climatic Change* 62: 189-216.
 1207 Xu, Z. and Z.L. Yang. 2012. An improved dynamical downscaling method with GCM bias corrections and
 1208 its validation with 30 years of climate simulations. *Journal of Climate* 25: 6271-6286.

1209 Xu, J., W. Long, J.D. Wiggert, L.W. Lanerolle, C.W. Brown, R. Murtugudde and R.R. Hood. 2012. Climate
1210 forcing and salinity variability in Chesapeake Bay, USA. *Estuaries and Coasts* 35: 237-261.
1211 Zeng, Z., W.W. Hsieh, A. Shabbar and W.R. Burrows. 2011. Seasonal prediction of winter extreme
1212 precipitation over Canada by support vector regression. *Hydrology and Earth System Sciences* 15:
1213 65-74.
1214 Zorita, E. and H. Von Storch. 1999. The analog method as a simple statistical downscaling technique:
1215 comparison with more complicated methods. *Journal of Climate* 12: 2474-2489.
1216
1217

Figure captions

Fig. 1: Study area with bathymetry for the Chesapeake Bay, Susquehanna River watershed, and major tributaries shown. The location of the Thomas Point buoy is shown in yellow, the location of eight weather stations providing air temperature observations are in red, and grid point locations for the WD GCM are shown in purple, to highlight the coarse spatial resolution of GCMs

Fig. 2: (a) Schematic representation of the statistical framework developed in this study. Models are boxed, model outputs are un-boxed, (b) conceptual model of a linear model tree (after Solomatine and Dulal 2003).

Fig. 3: Modeling daily surface water temperatures at the Thomas Point buoy: using daily air temperatures at Thomas Point (light gray), using a 17-day moving mean of air temperatures (dark gray), and using the non-linear equation from Mohseni et al. (2003) applied to a 17-day moving mean of air temperatures (black). The 1:1 ratio denoting perfect fit is shown as a black dashed line

Fig. 4: Water balance model results. Top: observed mean monthly Susquehanna River streamflow at Conowingo Dam (black), modeled monthly streamflow from the water balance model (red), and modeled monthly streamflow from the water balance model with the snow under-catch correction (green). Bottom: observed versus modeled monthly Susquehanna River streamflow 1970-200

Fig. 5: 2m air temperature and total precipitation anomalies for the Susquehanna River watershed from 26 GCMs under RCP8.5: 1956-2005 versus 2050-2099. The ensemble mean from all GCMs is shown in black, and extended to both axes with the black dashed line. The four GCMs chosen to represent the range of potential futures for the Chesapeake Bay are labeled

Fig. 6: Observed and modeled monthly surface temperature anomaly and surface salinity in the Upper and Lower Chesapeake Bay, and Potomac River, 1986 - 2015. Results from one well-sampled station in the Mid Bay are also shown (station CB5.1). Time series and R^2 statistics are shown for the dataset used to train the models (1986 – 2005), and the out-of-sample “test” dataset (2006 – 2015)

Fig. 7: Observed and modeled September surface temperature and salinity for a warm year (2008) versus a cool year (2011), and a wet year (2011) versus a dry year (2010). Results are interpolated (kriging) between CTD station locations (shown in black).

Fig. 8: (a),(b): Projections of mean Susquehanna River watershed 2m air temperature and total precipitation from 4 GCMs under RCP8.5, using four statistical downscaling methods. (c),(d): Ten year

moving means of air temperature and precipitation from each GCM, with the overall range from (a) and (b) shown in gray

Fig. 9: Ten year moving means of modeled (a) surface water temperature, (b) surface salinity and (c) Susquehanna River flow at Conowingo Dam from 4 GCMs under RCP8.5. A mean of the 4 statistical downscaling methods is shown, with overall range in gray. (a) and (b) are means across the entire Chesapeake Bay.

Fig. 10: Mean monthly surface temperature, Susquehanna River streamflow, and surface salinity for 1970-1999 (black), and 2071-2100 (red) under RCP8.5, from each GCM, averaged across all downscaling methods. Mean values for each time period are shown in bold lines, thin lines represent \pm one standard deviation among the thirty years in each time period

Fig. 11: a, c, e: Mean monthly surface temperature, Susquehanna River streamflow, and surface salinity for 1970-1999 (black), and 2071-2100 (colored) under RCP8.5 for the WW model only, for the BCQM (blue) and CFQM (pink) statistical downscaling methods. Mean values for each time period are shown in bold lines, thin lines represent \pm one standard deviation (omitted from future projections for clarity). b, d, f: As for a, c, e, but only future change (2071-2100 – 1970-1999) is shown, to highlight differences between downscaling methods.

Fig. 12: Projected changes in surface temperature in the Chesapeake Bay from the WW (a) and CW models (b) during summer: 2071-2100 minus 1970-1999, averaged across all downscaling methods. The two GCMs shown had the smallest and largest changes in temperature of the four considered. Results are shown on a common scale (4.1°C range, a, b), to highlight the difference between the two models, and on a model-specific scale (1.5°C range, c, d), to highlight spatial structure

Fig. 13: Projected changes in surface salinity in the Chesapeake Bay from the (a) WW and (b) WD models during winter: 2071-2100 minus 1970-1999. The two GCMs shown had the smallest and largest changes in salinity of the four considered

Figures intended as electronic supplementary material:

Fig. S1: Projected changes in surface temperature in the Chesapeake Bay during winter, spring, and fall: 1970-1999 versus 2071-2100. The two models with the strongest (top row) and weakest (bottom row) change between the two time periods are shown for each season

Fig. S2: Projected changes in surface salinity in the Chesapeake Bay during spring, summer, and fall: 1970-1999 versus 2071-2100. The two models with the highest (top row) and lowest (bottom row) change between the two time periods are shown for each season

Fig. S3: Mean 2m air temperature in the Susquehanna River watershed 2001-2100, from the four statistical downscaling methods applied to the WW model, averaged across seasons (top), and months (bottom). The difference between mean temperatures in August between the coolest (BCQM) and warmest (CFQM) methods is also shown

1284 **Table 1:** Predictor variables included in predictive model trees for surface temperature and salinity

	Surface temperature model	Surface salinity model
Predicted Thomas Point surface temperature (17d mean air temperature with non-linear correction)	X	
Longitude	X	X
Latitude	X	X
30d change in 17d air temperature	X	X
Previous 30d streamflow	X	X
Time of day	X	X

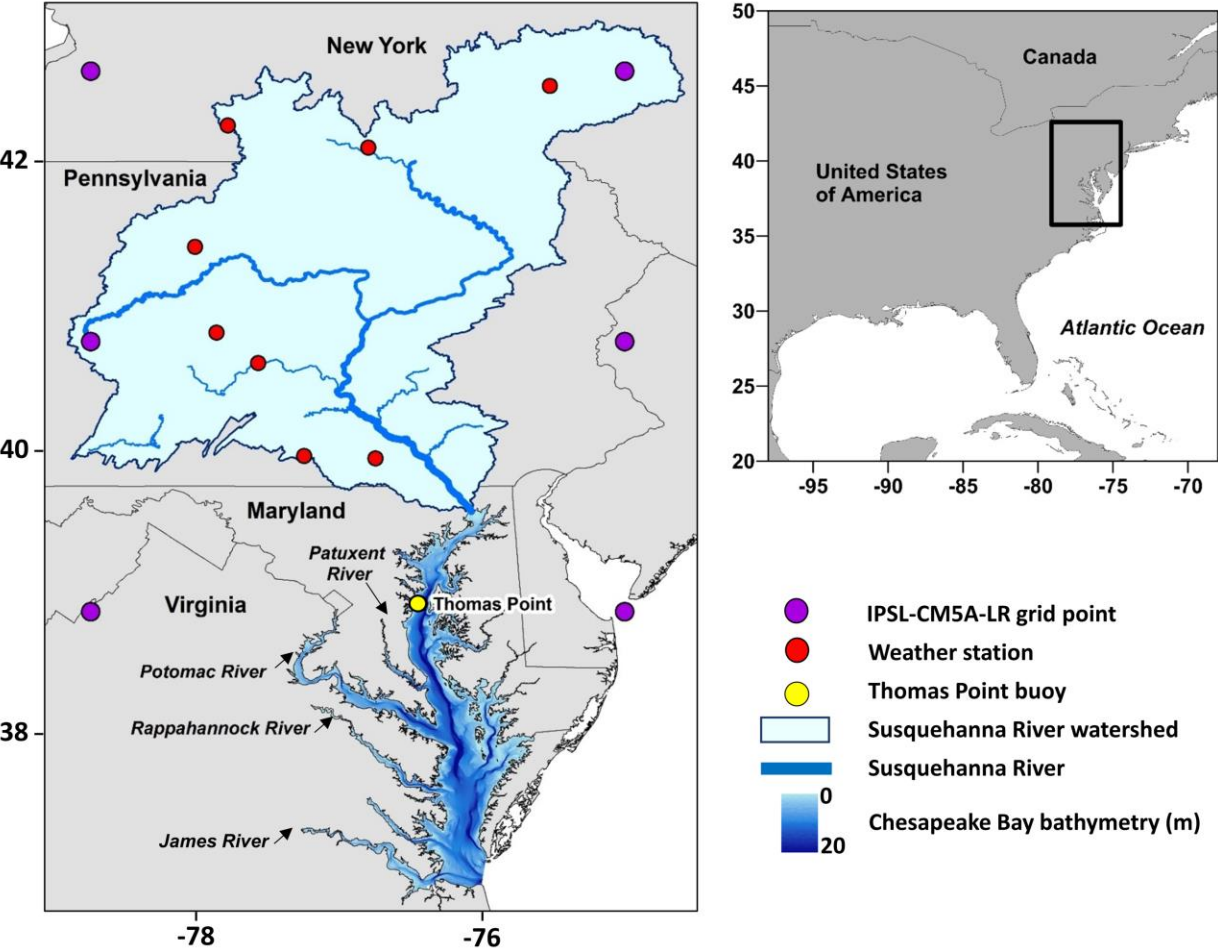
1285

1286

Table 2: R² values between observed and modeled surface temperature (monthly anomalies) and surface salinity in 7 regions of the Chesapeake Bay, from the unseen test data not used to build the predictive model trees. Mean annual surface temperature and salinity from CTD stations in each region are also shown, as are RMSE values between observed and modeled values.

	Mean surface temperature	Mean surface salinity	Surface temperature model: R ²	Surface salinity model: R ²	Surface temperature model: RMSE (°C)	Surface salinity model: RMSE
Lower Bay	17.80	17.96	0.74	0.55	1.14	2.43
Mid Bay	17.32	14.21	0.70	0.61	1.21	2.03
Upper Bay	17.70	9.01	0.73	0.76	1.31	1.80
James River	17.51	10.90	0.77	0.69	1.23	3.00
Rappahannock River	17.53	12.12	0.62	0.63	1.51	2.31
Potomac River	17.80	4.08	0.72	0.71	1.36	2.02
Patuxent River	17.60	10.75	0.63	0.42	1.32	2.19

1298 Fig. 1



1299

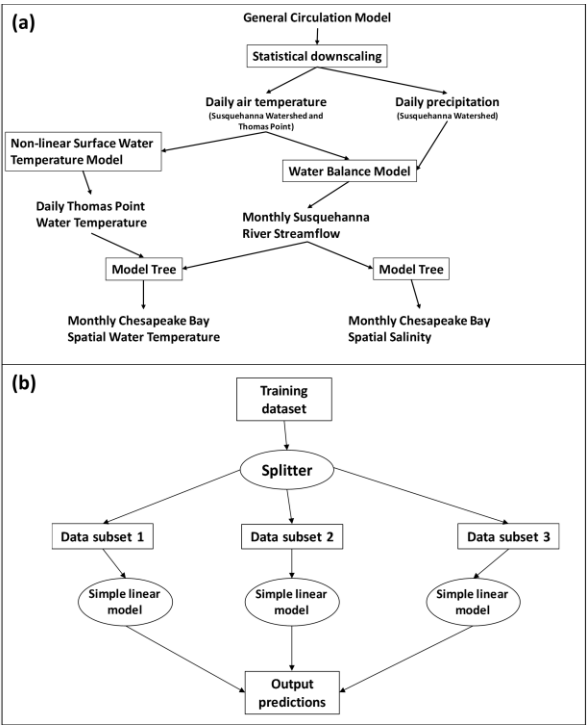
1300

1301

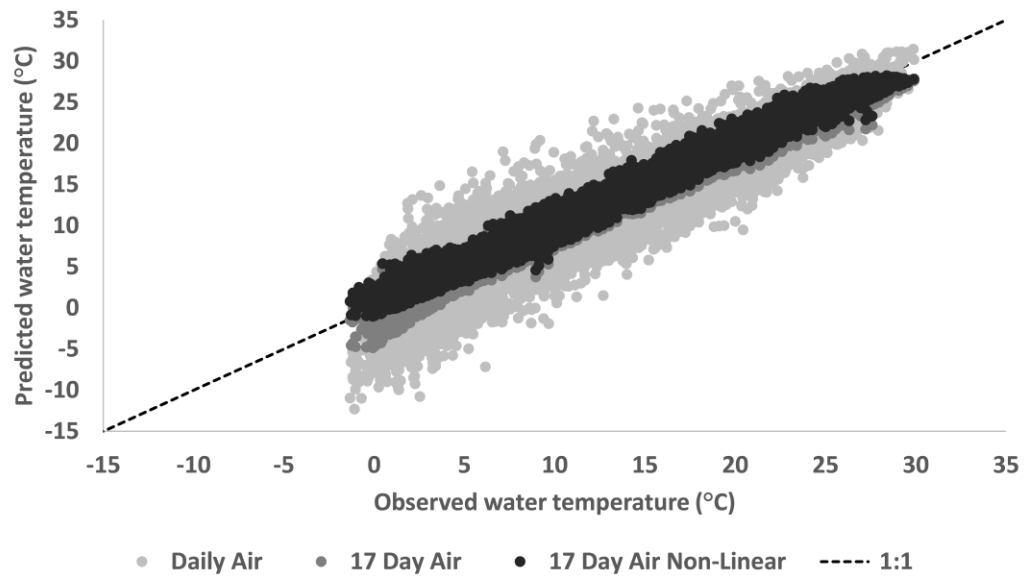
1302

46

1303 **Fig. 2**



1309 **Fig. 3**



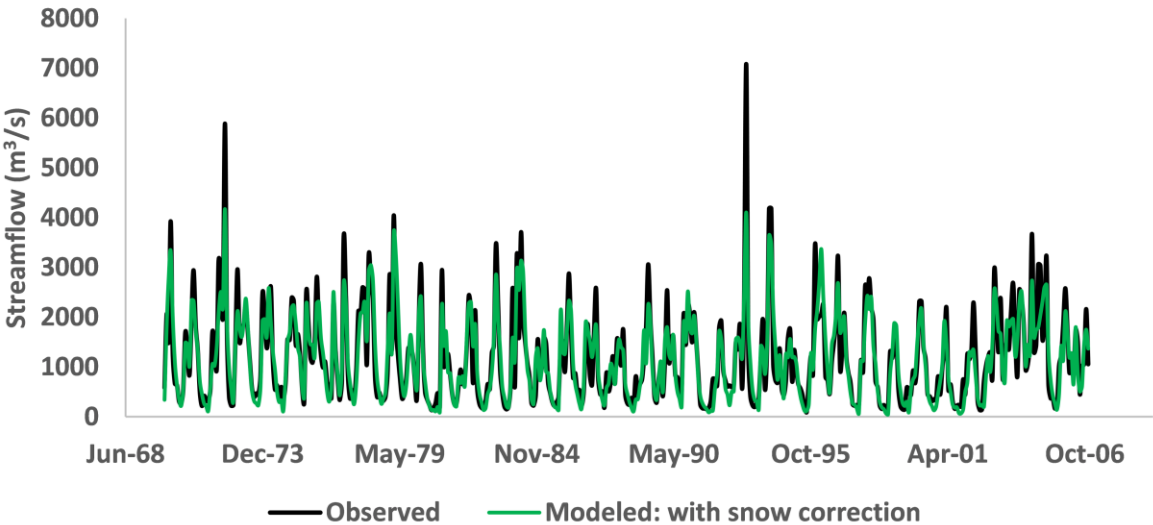
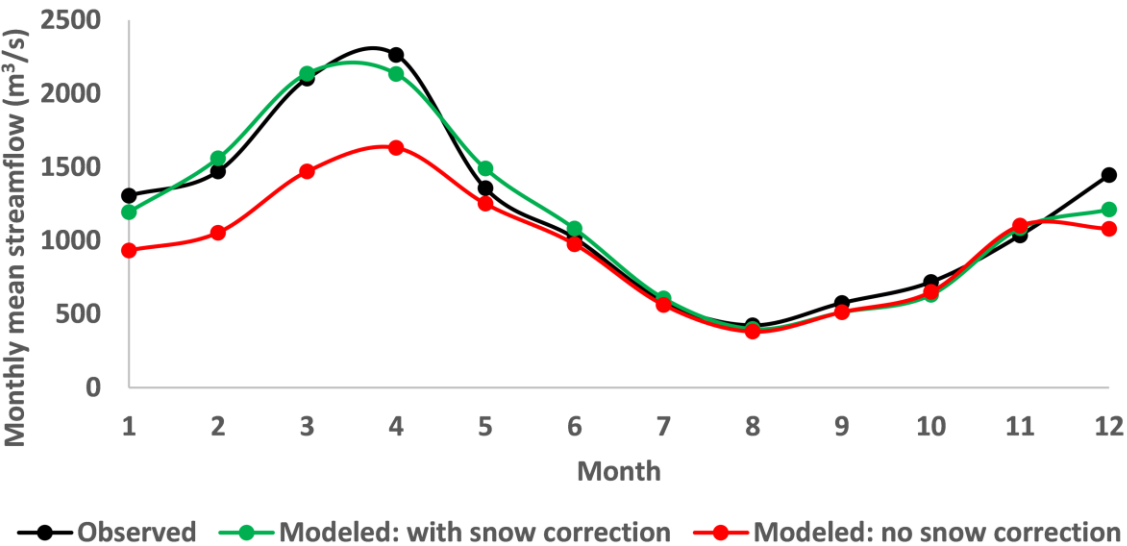
1310

1311

1312

1313

1314 **Fig. 4**

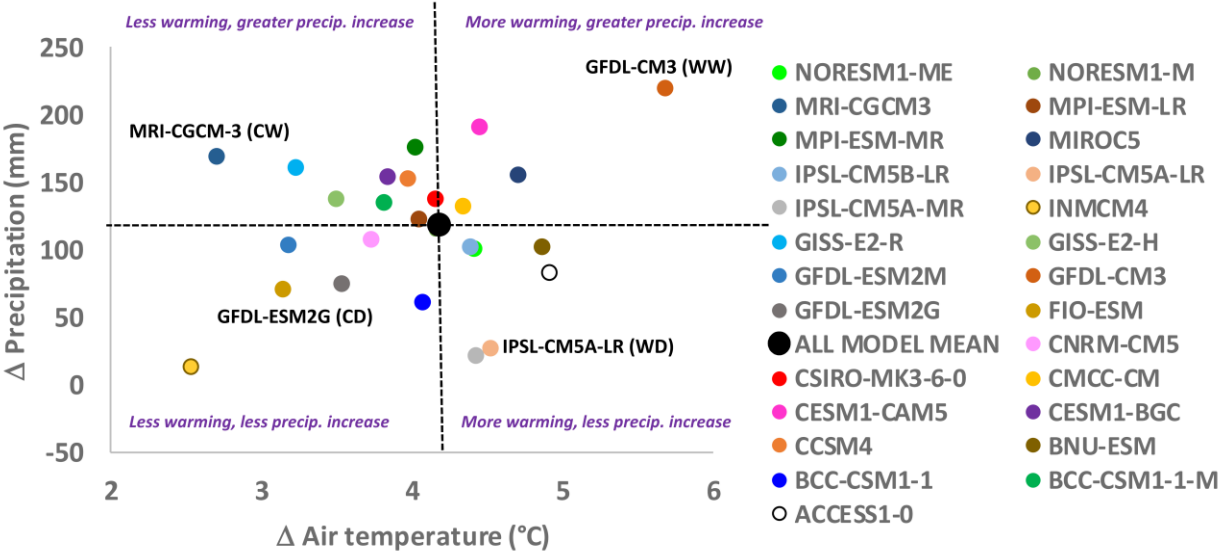


1315

1316

1317

1318 Fig. 5

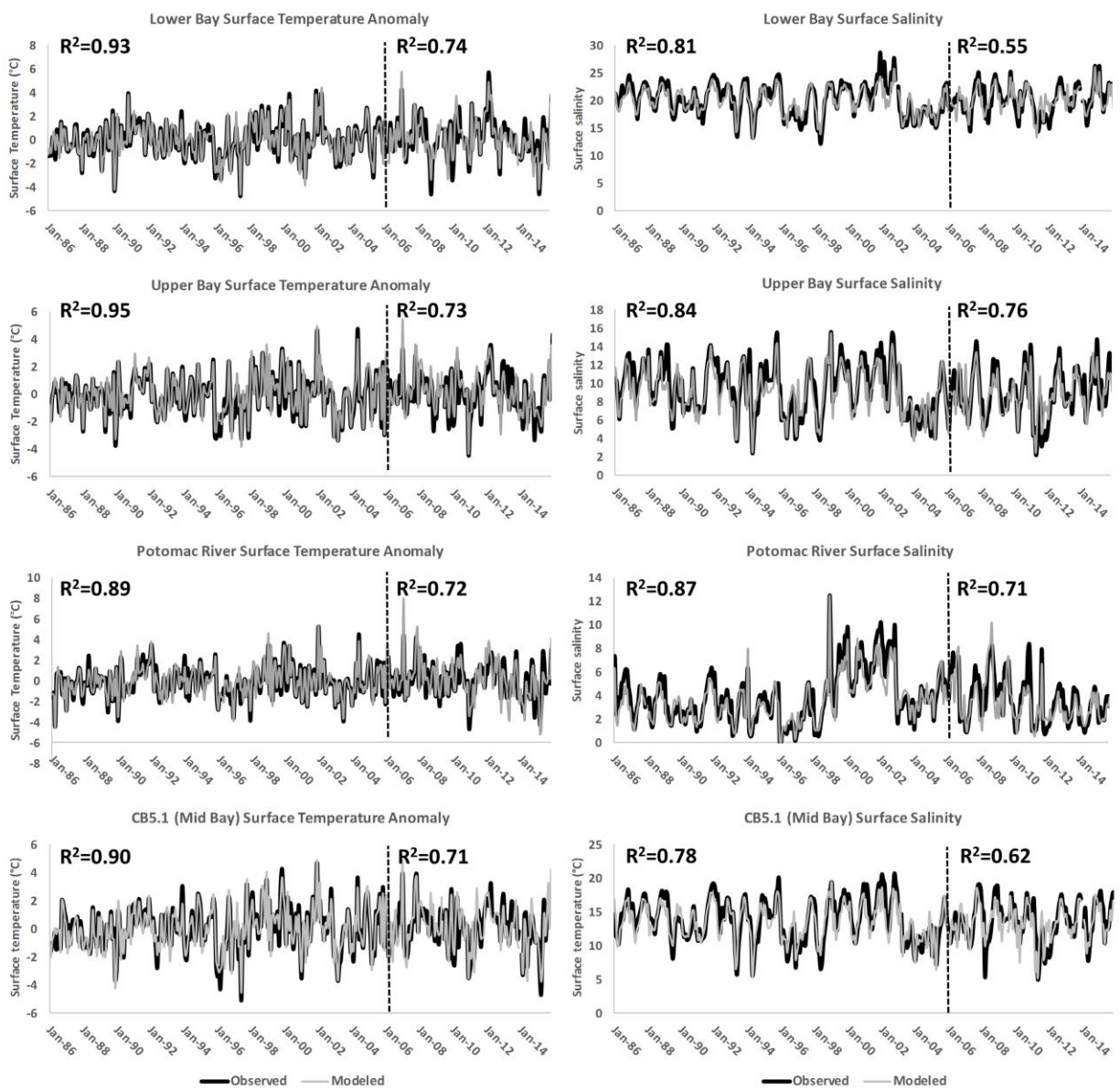


1319

1320

1321

1322 **Fig. 6**

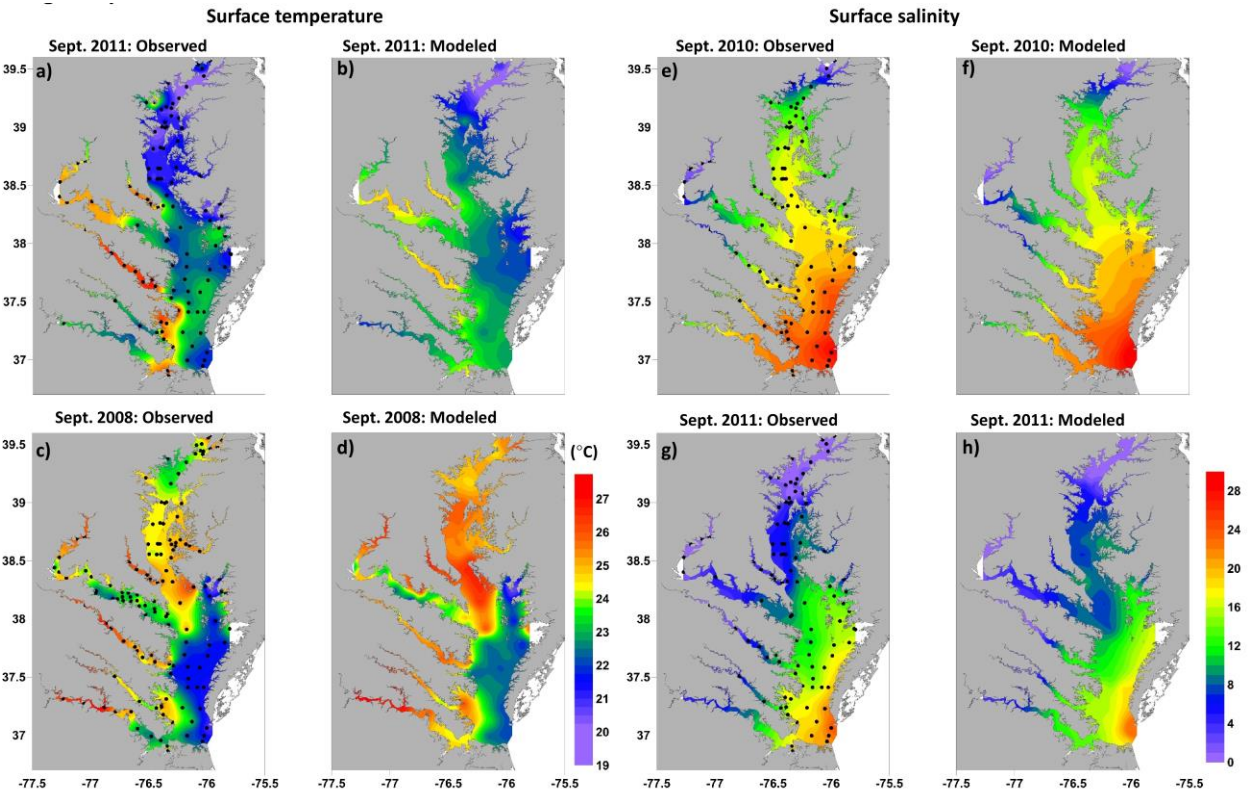


1323

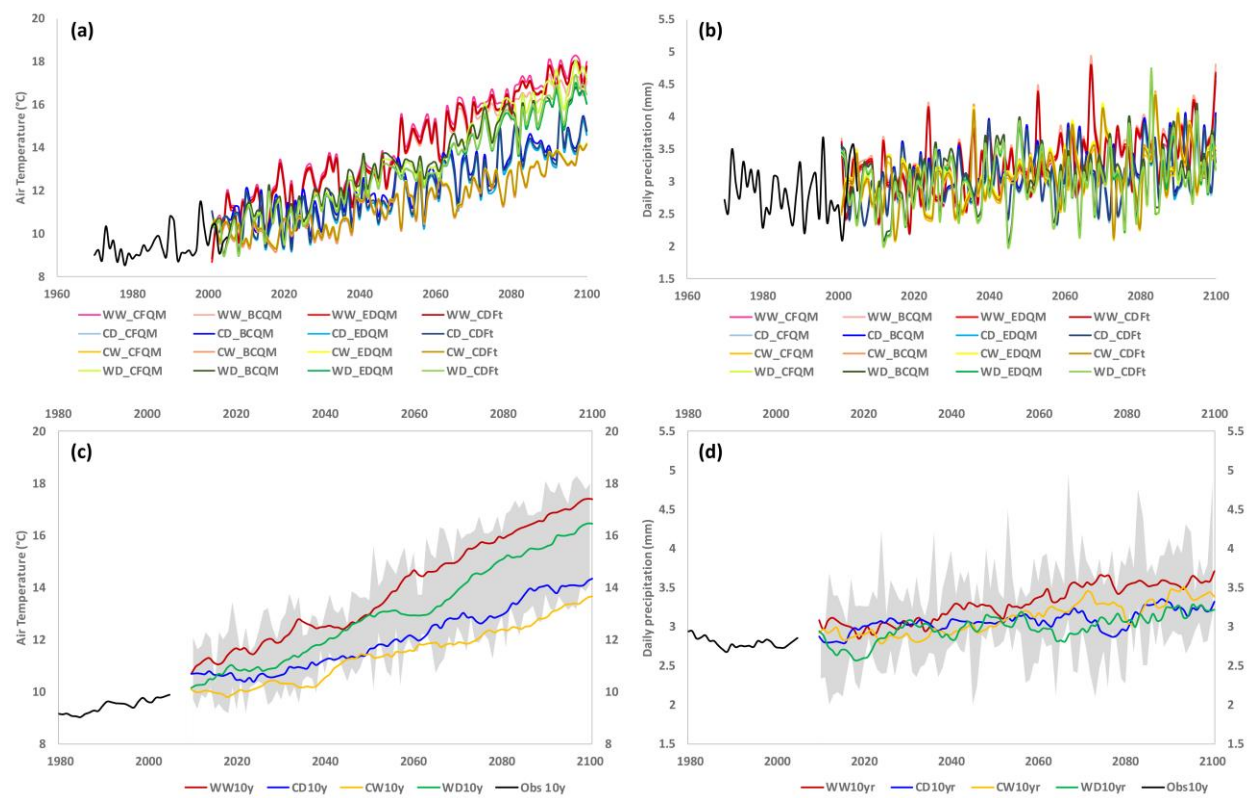
1324

1325

1326 **Fig. 7**



1330 **Fig. 8**

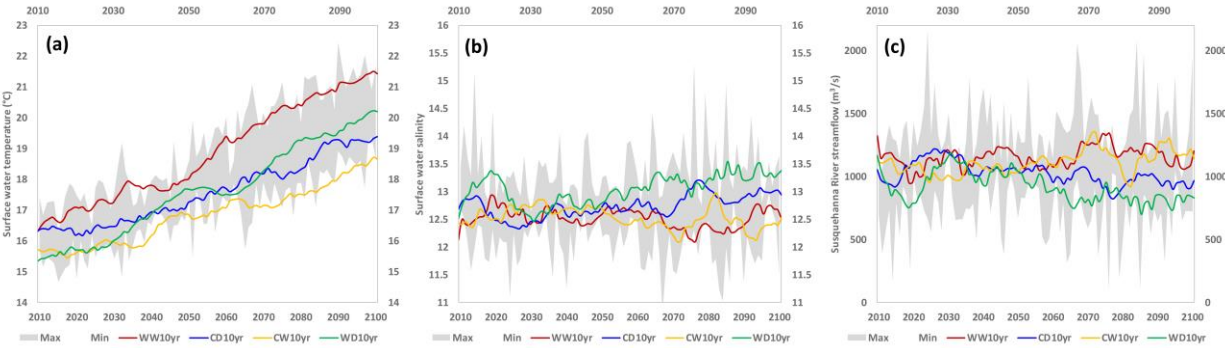


1331

1332

1333

1334 **Fig. 9**

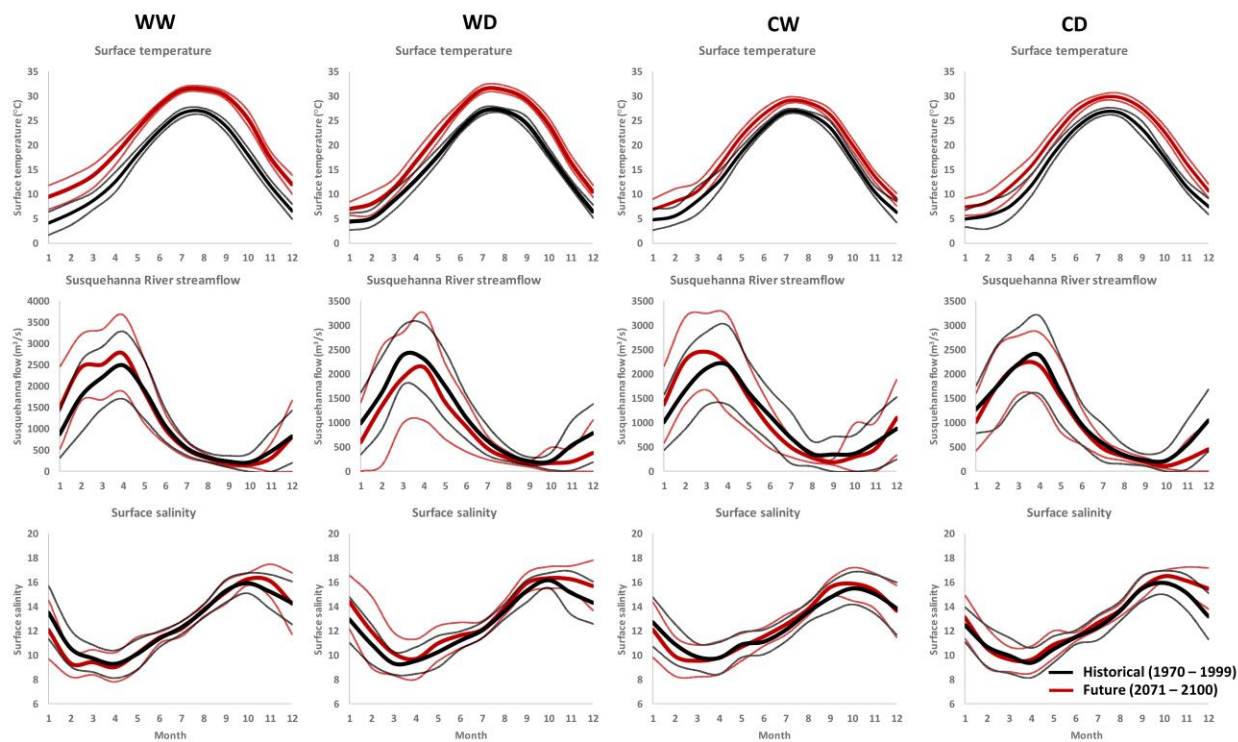


1335

1336

1337

1338 **Fig. 10**

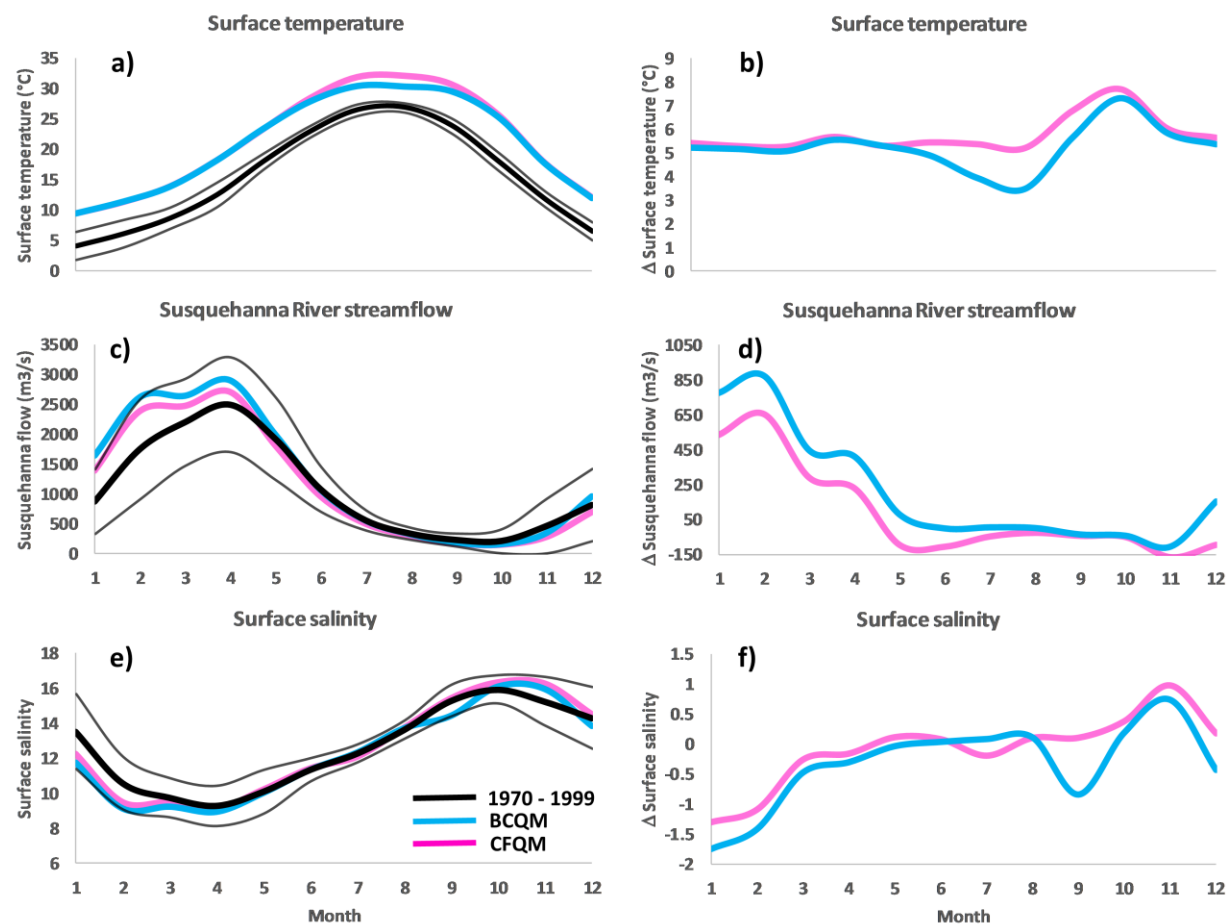


1339

1340

1341

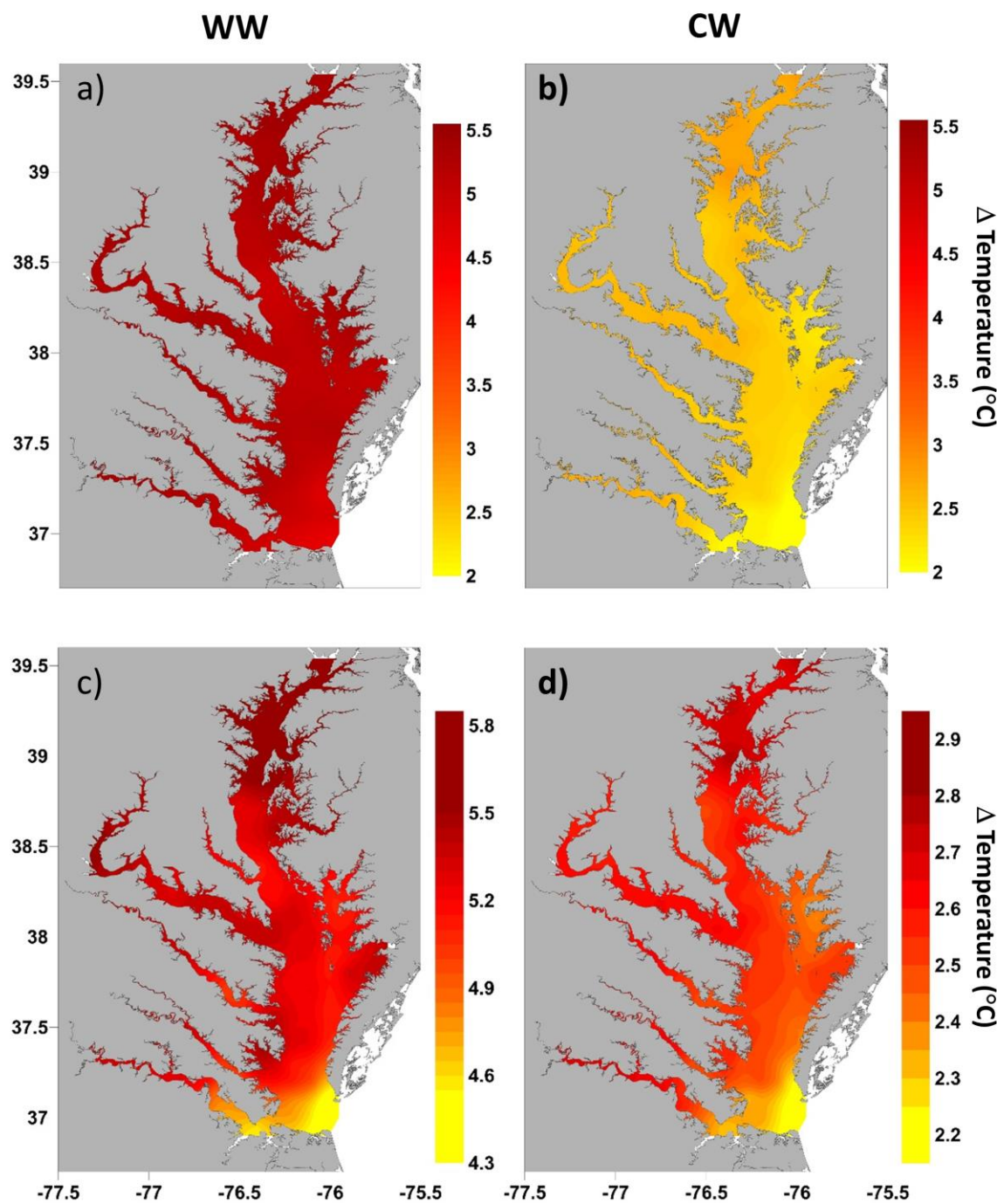
1342 **Fig. 11**



1343

1344

1345

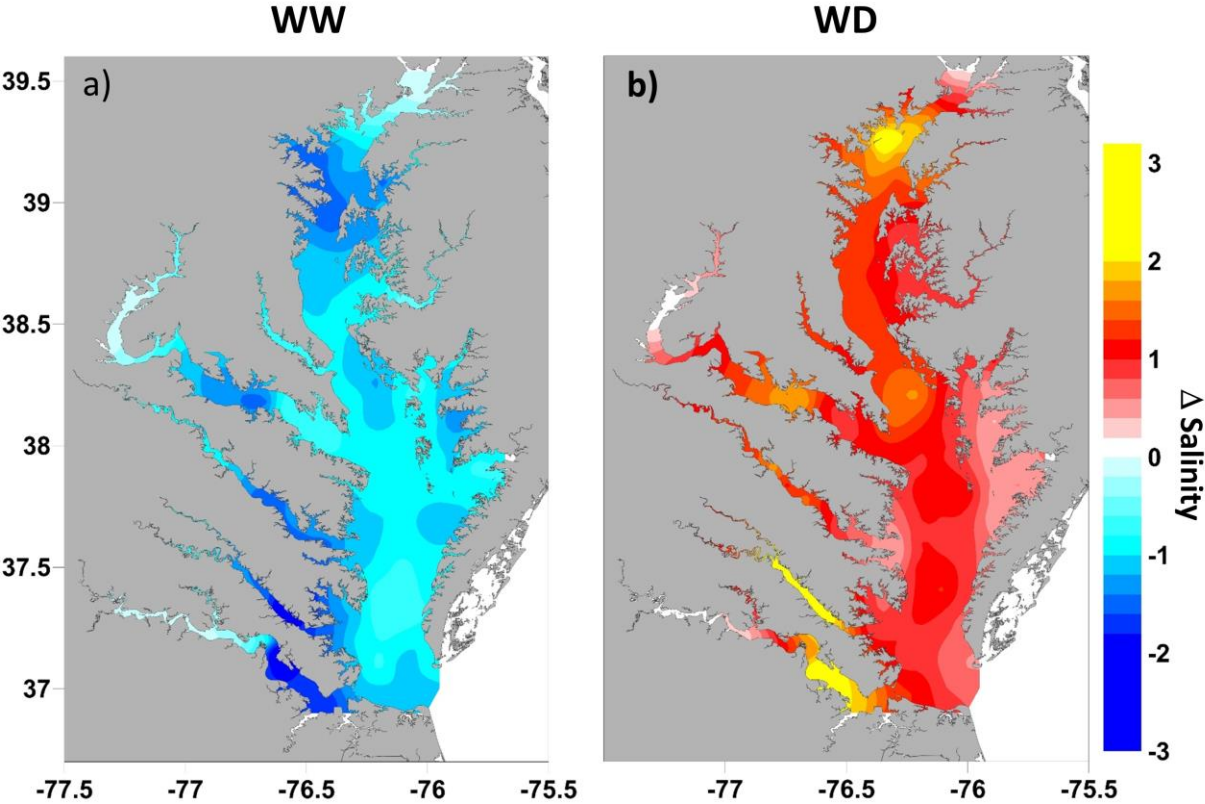


1347

1348

1349

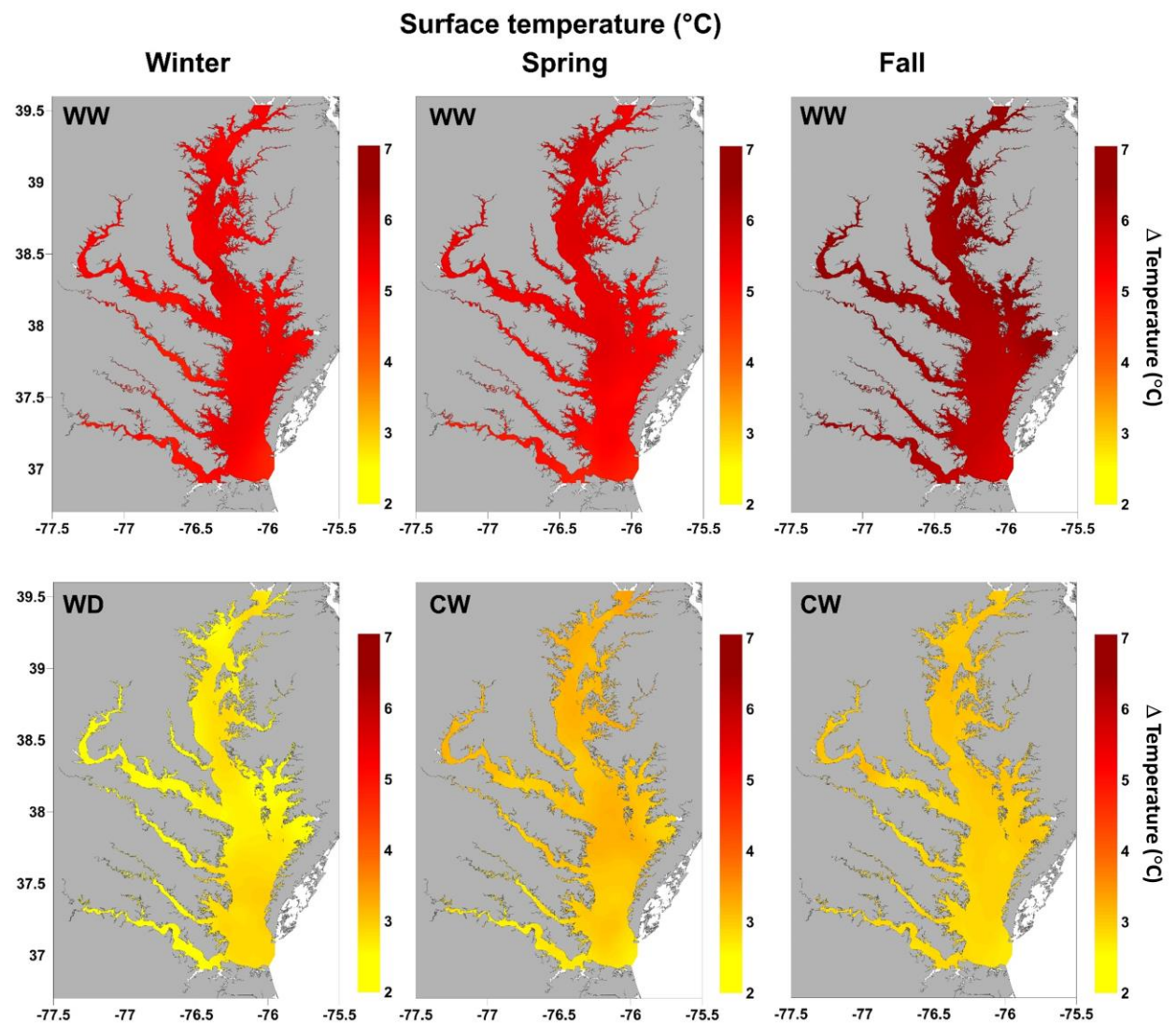
1350 **Fig. 13**



1351

1352

1353 **Fig. S1**



1354

1355

1356

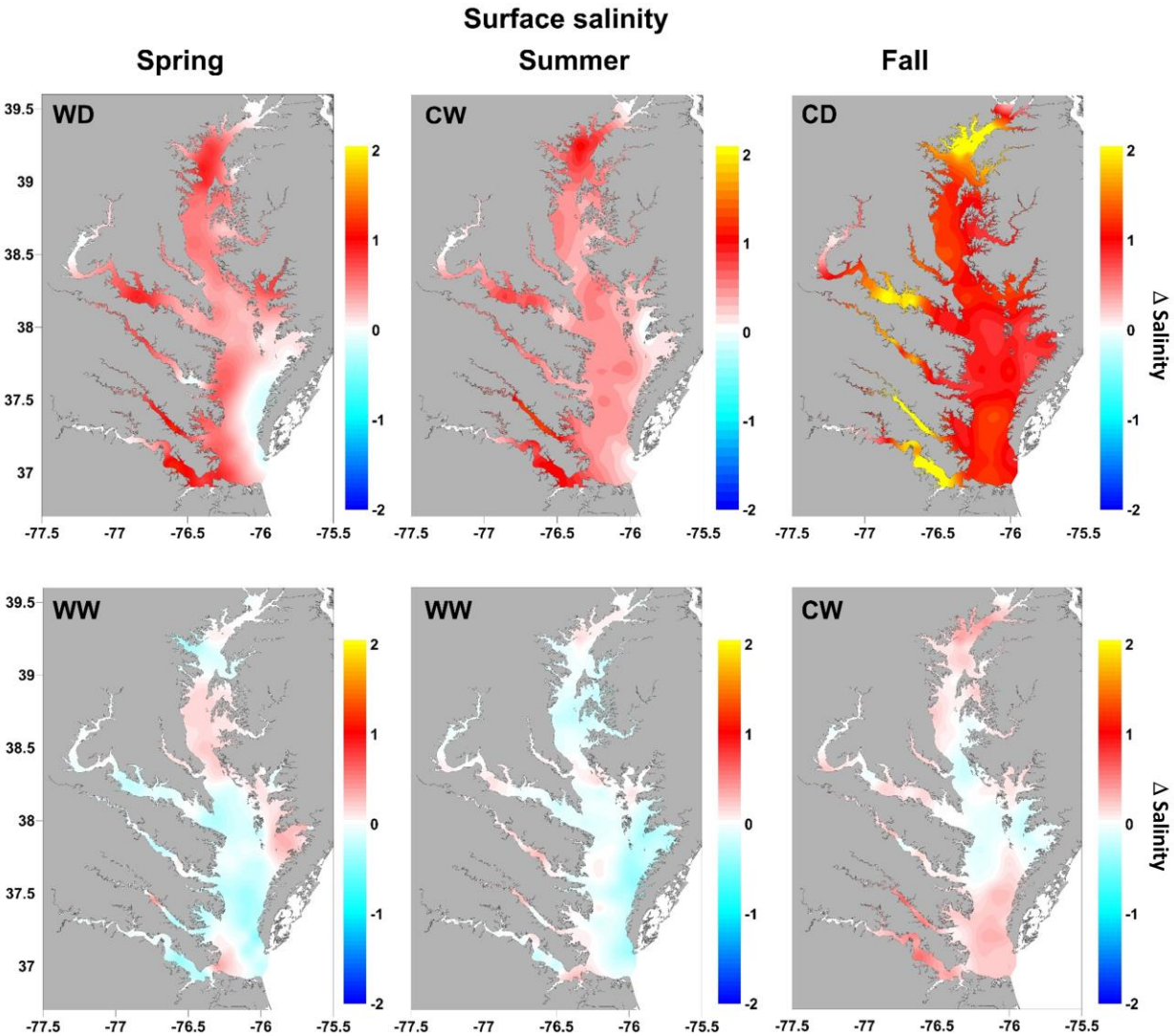
1357

1358

1359

1360

1361



1363

1364

1365

1366

